International Institute of Humanitarian Law

International Institute of Humanitarian Law
Institut International de Droit Humanitaire
Istituto Internazionale di Diritto Umanitario

Whither the Human in Armed Conflict? IHL Implications of New Technology in Warfare

STUDI

Politica

FrancoAngeli

International Institute of Humanitarian Law
Institut International de Droit Humanitaire
Istituto Internazionale di Diritto Umanitario

Whither the Human in Armed Conflict? IHL Implications of New Technology in Warfare

42nd Round Table on Current Issues
of International Humanitarian Law
(Sanremo, 4th-6th September 2019)

Editor Gabriella Venturini

Associated Editor Gian Luca Beruto

FrancoAngeli

Gabriella Venturini is Professor Emerita of International Law, University of Milan, Italy and Council Member, IIHL. She is the President of the Italian Branch of the International Law Association. She taught and wrote extensively in the field of public international law and EU law and collaborated with the Italian Ministry of Foreign Affairs as a member of the Italian delegation at several international meetings, conferences and negotiations.

Gian Luca Beruto holds a Master's degree in International Political Science and is currently Assistant to the Secretary-General of the International Institute of Humanitarian Law. His career has developed in different areas including peacekeeping, conflict management, institutions building, governance, disarmament, assistance and protection of migrants and refugees throughout Europe and Africa.

The International Institute of Humanitarian Law would like to thank
Ms Shirley Morren, librarian of the Institute and Mr. Edoardo Gimigliano,
who were involved in the painstaking task of proofreading and editing.

Stampa: Geca Industrie Grafiche, Via Monferrato 54, 20098 San Giuliano Milanese.

Table of Contents

Preface

The tight bond linking scientific progress, technological development and their military exploitation constitutes a recurring leitmotif in the current international security scenario.

In this context, the rule of law may appear blurred and international humanitarian law (IHL) incapable of keeping pace with technological progress. Armaments research and innovation programmes are among the issues at the top of the agenda of advanced national powers, bolstering an already unbalanced relationship between growing technological military capacity and the legal frameworks which limit their usage according to IHL fundamental principles.

The 42nd Round Table on current issues of international humanitarian law, jointly organized by the Sanremo Institute and the International Committee of the Red Cross, gathered together academics, legal experts, military commanders and government officials to discuss the crucial question of technology developments and their application in armed conflicts, including the challenges imposed by the military use of cyber arms and the widespread of autonomous weapons in warfare, focusing on the risks related to their use in urban contexts.

The Round Table provided the opportunity for fruitful and constructive debates on crucial topics, such as the potential human costs of cyber warfare and applicable IHL provisions; the potential support that cyber technology could provide to humanitarian operations; and how IHL represents an effective legal framework for warfare in outer space. The way forward in addressing the challenges of using new technologies and weapons within military operations were also discussed, highlighting the primary objective of IHL to protect and safeguard civilians and vulnerable groups from the violence of armed conflicts.

The proceedings of this Round Table aim to confirm, once again, the "humanitarian dialogue in the spirit of Sanremo" and to strongly reassert the importance of promoting the application of IHL, particularly when it comes to specific areas where regulatory gaps occur.

Edoardo GREPPI
President of the International Institute of Humanitarian Law

Opening session

Welcome address

Alberto BIANCHERI
Mayor of Sanremo

Sono particolarmente onorato di porgere, a nome di tutta l'Amministrazione Comunale, il più caloroso benvenuto a tutte le personalità che prendono parte a questa Tavola Rotonda sui problemi attuali del diritto internazionale umanitario, organizzata congiuntamente dall'Istituto Internazionale di Diritto Umanitario di Sanremo e dal Comitato Internazionale della Croce Rossa di Ginevra, che giunge quest'anno alla 42ª edizione.

Vorrei limitarmi a qualche parola per sottolineare la mia grande soddisfazione e il sincero orgoglio che ho nel rappresentare la città in cui ha sede questo prestigioso Istituto – di cui il Comune di Sanremo è cofondatore – che da quasi 50 anni lavora assiduamente per promuovere in tutto il mondo il rispetto del diritto internazionale umanitario e dei diritti umani.

L'Istituto, grazie al suo prestigio sul piano internazionale costituisce, non solo per la città di Sanremo ma per il Ponente Ligure e tutta la Regione, una importante risorsa il cui operato ha tangibili e positivi risultati sul territorio.

La Tavola Rotonda, organizzata ogni anno nel mese di settembre – e che si pregia della Targa del Presidente della Repubblica Italiana e del patrocinio del Ministero degli Affari Esteri e della Cooperazione Internazionale e del Ministero della Difesa – rappresenta un consolidato appuntamento internazionale, apprezzato in tutto il mondo che approfondisce le problematiche umanitarie di maggiore attualità.

Il tema affrontato quest'anno è particolarmente interessante. Gli attacchi "cyber", il sempre più frequente utilizzo dell'intelligenza artificiale nei sistemi di armamento, la "guerra spaziale", le problematiche derivanti dall'uso di tali tecnologie in teatri di guerra e dallo sviluppo tecnologico applicato alle operazioni militari, con evidenti rischi per la popolazione civile, saranno tra le principali questioni esaminate nel corso dei lavori con l'obiettivo di chiarire se è necessario mantenere, oppure no, un controllo umano diretto sulle nuove tecnologie che operano nei conflitti presenti e futuri.

È evidente come, nell'odierna realtà internazionale caratterizzata da continue violazioni che colpiscono profondamente e sistematicamente

l'integrità, la dignità e la sopravvivenza delle fasce più vulnerabili della popolazione civile, lo sviluppo tecnologico sia necessariamente un fenomeno da controllare, proprio per limitare tali violazioni.

Sono più che mai convinto che con il contributo di autorevoli rappresentanti di Governi e delle principali Organizzazioni Internazionali, di eminenti studiosi ed esperti provenienti dalle diverse aree geografiche del mondo, la Tavola Rotonda di Sanremo sarà, ancora una volta, l'occasione per un costruttivo scambio di punti di vista e di esperienze tra tutte le parti interessate.

Sono particolarmente lieto, anche a nome di tutta la cittadinanza, di poter esprimere ai presenti il mio augurio di buon lavoro con il più sincero auspicio che nel corso di questo breve soggiorno potrete trovare anche il tempo per scoprire le bellezze e le attrattive che offre questa città.

Spero di rivedervi presto a Sanremo.

Opening remarks

Fausto POCAR
President, International Institute of Humanitarian Law (IIHL)

Excellences, Autorités civiles et militaires, estimé(e)s collègues et cher(e)s Ami(e)s, Mesdames et Messieurs,

C'est pour moi un privilège et un grand honneur d'ouvrir encore une fois la traditionnelle table ronde de l'Institut, parvenue à sa 42ème édition, consacrée à l'examen de problèmes actuels du droit international humanitaire. Et je me réjouis que la table ronde est organisée, comme d'habitude, en coopération avec la Comité International de la Croix-Rouge, renouvelant ainsi une synergie fructueuse qui a permis de réunir un nombre important d'experts – à la fois académiques, militaires et opérateurs sur le terrain – dans l'environnement neutre et amical de notre Institut qu'on appelle depuis longtemps "l'esprit de Sanremo", pour en souligner la qualité des débats qui y ont lieu dans ses nombreux cours et dans la table ronde.

C'est bien dans cet esprit que j'adresse à tous et à toutes les personnes qui se trouvent dans cette salle la bienvenue la plus chaleureuse au nom de l'Institut et mon personnel ma vive gratitude pour s'être rendu(e)s à Sanremo pour participer à cette Table Ronde, dédiée à un sujet tout particulièrement actuel tel que les implications du recours à de nouvelles technologies dans la conduite des conflits armés.

Nel porgere il mio saluto ai partecipanti, desidero esprimere la mia profonda gratitudine al Presidente della Repubblica, che ha voluto sottolineare ancora una volta il suo apprezzamento per questo evento conferendo alla Tavola Rotonda la "Medaglia del Presidente della Repubblica". È un alto riconoscimento che onora l'Istituto e ci incoraggia a continuare con sempre maggiore impegno la nostra attività di insegnamento e di dibattito intesa ad assicurare un maggiore rispetto del diritto internazionale umanitario.

Vorrei anche esprimere il mio ringraziamento alle autorità civili e militari presenti in sala e alle illustri personalità che prenderanno la parola in questa sessione di apertura della Tavola Rotonda: al Sindaco di Sanremo, Alberto Biancheri, recentemente rieletto per un secondo mandato e qui rappresentato da Alessandro Sindoni, che ringrazio unitamente a tutti i componenti della giunta comunale per il costante sostegno a favore

dell'Istituto; al Presidente del CICR, Ambasciatore Peter Maurer, che pur non potendo essere presente ha inviato un messaggio con un video; alla Direttrice per il diritto e la politica internazionale del CICR, Helen Durham, che come negli anni passati ha dato un contributo importante alla preparazione della tavola rotonda; e al Direttore generale degli affari politici e di sicurezza del Ministero per gli affari esteri e la cooperazione internazionale, Ambasciatore Sebastiano Cardi, che ha assicurato l'appoggio del Ministero al nostro incontro e che, pur non potendo intervenire per impegni a Roma, ha inviato un messaggio scritto che sarà letto e figurerà negli atti della tavola rotonda.

Ainsi que le programme l'indique, cette Table Ronde a lieu à l'occasion du 70ème anniversaire des Conventions de Genève du 12 avril 1949. Le colloque est toutefois bien loin d'être une simple célébration de cet anniversaire qui concerne des conventions qui sont très connues et dont l'importance ne saurait être mise en doute. Il s'agit en effet non seulement de tracer un bilan de la contribution que ces conventions ont donné à l'évolution du droit international humanitaire et à l'affirmation des principes desquels il s'inspire, mais également d'identifier le rôle qu'elles peuvent jouer dans le cadre de la complexité des scénarios des conflits armés qui se déroulent actuellement dans le monde: un contexte nouveau qui est, entre autre, caractérisé par l'emploi de nouveaux moyens et de nouvelles méthodes de combat, largement favorisé par de nouvelles technologies, ainsi qu'on l'a déjà souligné aux cours des débats de la 39ème table ronde, qui a eu lieu en 2016, de laquelle la présente constitue d'une certaine manière la continuation.

In light of the above-mentioned considerations, this Round Table will focus on the impact of new technologies on IHL, on how IHL responds to technological development, and on the role of the human in a scenario where new technologies increasingly assist or even replace the human in warfare. Of course, the question of dealing with technological development is not new for IHL. New weapons have been invented continuously, in peacetime and in wartime, as for example dramatically happened with the development of nuclear weapons during WWII. Incidentally, this is also an anniversary, as WWII started exactly at the beginning of September 80 years ago. Whatever the technological changes, however, and whatever the adaptations that may be required in the law, it should remain clear that the basic principles of IHL continue to apply to existing and future means and methods of warfare, as they are enshrined in the Geneva Conventions of

1949 and the Additional Protocols of 1977. Nevertheless, it is important to discuss how changes in warfare have been considered by IHL so far, in order to draw lessons for future approaches.

The Round Table will address some of the most debated issues concerning new technology in warfare. Therefore, it will deal with the challenges raised by cyber technology in contemporary armed conflicts, bearing in mind that many States already have, or are in the process of building, offensive capabilities in the cyber space. Another sensitive topic is the use, by the military of an increasing number of States, of autonomous weapons, capable of autonomously selecting military targets.

The role of the human to control such weapons is a matter for debate not only from the ethical point of view, but also in light of the principles of accountability and responsibility.

New technologies will also affect the possibility of future outer space warfare. The militarisation of outer space was largely debated in the second half of last century within the UN, with a view to preventing it in the interest of humankind. While the debate appeared less significant in the aftermath of the end of the cold war, it has been recently revitalized and the challenges that IHL may face, should an armed conflict occur in outer space, deserve to be explored and discussed.

From outer space back to Earth, the Round Table will address specific features of contemporary armed conflicts, which frequently imply military operations in an urban context. How is new technology used in such a context, and how it may assist in observing fundamental principles of IHL, in particular the principles of distinction and precaution, is a matter for debate, bearing in mind that in that context the civilian population is especially exposed.

The panels of the Round Table will endeavour to address the challenges that I have very succinctly mentioned and additional ones, with a view to clarifying the factual and legal framework of contemporary armed conflicts as far as the use of new technology in warfare is concerned. The qualifications and experience of the speakers and the moderators, as well as the participants who are attending this Round Table, will no doubt ensure interesting and lively debates, in line with the tradition of our most successful round tables.

Let me conclude by anticipating that the foreseeable success of this event will also constitute an excellent introduction to the forthcoming 50th anniversary of our Institute, which will be celebrated next year, in 2020.

Opening remarks

Helen DURHAM
Director of International Law and Policy,
International Committee of the Red Cross (ICRC)

Excellences, ladies and gentlemen, dear friends and colleagues,

It is my pleasure to join Professor Pocar in welcoming you to the 42nd Sanremo Round Table.

This year's Round Table takes place on the 70th anniversary of the four Geneva Conventions of 1949. Thus, before I turn to contemporary and future challenges related to new technologies in warfare, I would like to suggest three lessons from the remarkable success of the Geneva Conventions that might be useful for our discussions in the coming days.

Firstly, defining limits for warfare is possible. In 1949, only four years after the tremendous suffering of the Second World War, and at the beginning of the Cold War, States negotiated the 429 Articles of the Geneva Conventions in only 4 months. Since then, the Conventions have achieved universal ratification. To us, this is a compelling example of what can be achieved when States come together, driven by the common purpose to preserve a minimum of humanity even in times of armed conflict.

Secondly, international humanitarian law has a real impact on armed conflicts. When we look at the news, read the reports of fact-finding missions, or follow international criminal law trials, we see shocking and unacceptably high levels of suffering caused by armed conflicts and by a lack of respect for IHL. In light of this reality, some may ask whether IHL is still relevant, and whether it is worth thinking about new rules. I am not convinced by narratives on the 'erosion' of IHL. While the ICRC witnesses the horrors of armed conflict firsthand, it is in exactly these conflicts that we also see how IHL is respected! We see quiet, everyday achievements of IHL – when a military takes care in its targeting to not fire on civilian buildings; when a wounded person is allowed through a checkpoint; when a child on the frontlines receives food and other humanitarian air; and when detainees are able to send a message to their families. These success stories prove that respect for IHL is possible and happening.

And thirdly, we do not have to reinvent the wheel. The Geneva Conventions and other rules of IHL remain today as relevant as 70 or 40 years ago: IHL is up to the contemporary challenges. IHL does not ask the impossible. States were not carried away by idealism when they negotiated

the Geneva Conventions. They designed a body of law for extreme circumstances of armed conflict, striking a careful, pragmatic balance between military necessity and humanity.

And most importantly, when States adopt IHL treaties, they always do so with a view to regulating future conflicts – which are likely to involve means and methods of warfare unknown at the time the treaty is negotiated.

These positive lessons should, however, not divert our attention from the fact that more needs to be done. Respect for IHL is far from perfect and parties to armed conflicts need to invest more into implementing IHL. Moreover, existing rules of IHL are not always clear – we need government and non-governmental lawyers to interpret and clarify the law. One important occasion to do so is at this Sanremo Round Table on the IHL implications of new technologies of warfare.

When we discuss about new technologies in the next two days – such as cyber tools, autonomous weapons systems, artificial intelligence, or weapons in outer space – we will not speak about hypothetical or abstract future developments. In fact, some of these means or methods of warfare are used in contemporary armed conflicts, and an increasing number of States are developing relevant capacities.

In the ICRC's view, new technologies in warfare hold great promise, but they are posing great risks. Technological advances can certainly have positive effects on the protection of civilians in armed conflict when used to that end: weapons can be used with more precision; military decisions can be better informed; and military aims can be achieved without the use of kinetic force or physical destruction. At the same time, new means of warfare and the way they are employed can also pose new risks to combatants and civilians. Moreover, some of them pose questions on the role of the human in warfare.

For example, cyber operations during the past years have shown the potential human cost that cyber operations can cause. Attacks on the medical sector, attacks on critical infrastructure, or attacks on plants containing dangerous forces can cause significant human harm. While today's cyber operations have fortunately not lead to human casualties, much is unknown on how this technology will evolve, which capabilities and tools the most sophisticated actors develop, and to what extent the use of cyber operations during armed conflicts might be different from the trends observed so far.

The development of autonomous weapons system, including systems that incorporate artificial intelligence and machine learning, pose their own set of issues. For the ICRC, the primary concern is a loss of human control

over the use of force. If the user of an autonomous weapon system is uncertain about the exact timing, location and circumstances of the actual use of force, the effects are difficult to predict. This poses important risks for civilians in the area where this weapon system is used.

When we think about new technologies of warfare from a legal point of view, the ICRC is of the firm view that IHL applies to the use of new means and methods of warfare. As I said before, IHL treaties are necessarily developed with regard to future conflicts, and rules such as those on legal reviews of new weapons underline that States consider existing rules relevant and applicable to future means and methods of warfare. On this issue, we are closely aligned with many States and with the International Court of Justice, which has stressed that the established principles and rules of humanitarian law apply 'to all forms of warfare and to all kinds of weapons', including 'those of the past, those of the present and those of the future'[1]. Importantly, however, acknowledging IHL applicability does not legitimize the use of new technologies of warfare. Moreover, it does not set aside *ius ad bellum* – the UN Charter must be respected in all circumstances. IHL defines additional limits if parties to armed conflicts decide to employ new technologies in warfare.

The recognition that IHL applies to the use of new technologies in warfare brings us to what we believe should be the core of discussion among States and other experts, namely the question of how IHL applies to the digitalization of warfare.

Weapons that can select and attack targets without human intervention have been a focus of discussions among States for the last six years. But even when discussing weapons that can operate autonomously, we must always keep in mind that it is humans that have to comply with and implement IHL. This responsibility cannot be transferred to a machine or a computer program. In the ICRC's view, human control must be maintained for both legal and ethical reasons. Combatants need to retain a level of control that allows them to make the context-specific legal judgements in specific attacks as required, for instance, by the IHL rules on distinction, proportionality and precautions in attack. The design or use of an autonomous weapon should not prevent the user from making these judgements. In the ICRC's view, legal and ethical concerns should inform the establishment of internationally agreed limits on autonomy in weapon

[1] International Court of Justice, *Threat or Use of Nuclear Weapons*, Advisory Opinion, 8 July 1996, para. 86.

systems. A human-centered approach will also be necessary to ensure that any broader applications of artificial intelligence and machine learning in armed conflict – such as in decision-making – preserve the necessary human judgement.

With regard to cyber warfare, our key message is that despite the interconnectivity that characterizes cyber space, the principles of distinction, proportionality and precautions can be respected. Cyber tools are not inherently indiscriminate. In fact, many of the cyber attacks that have been observed appear to have been rather discriminate from a technical perspective. Being able to target an attack will not necessarily make the attack lawful under IHL – but it shows that fundamental IHL principles on the conduct of hostilities can be respected.

However, IHL rules protecting civilian objects may only provide the full scope of legal protection if States recognize that cyber operations impairing the functionality of civilian infrastructure are subject to the rules governing *attacks* under IHL. Moreover, the 'datafication' of our societies makes it necessary to recognize that civilian data are afforded the same protection as physical civilian objects: it simply does not make sense to accept that traditional archives qualify as civilian objects and are protected against attack but that digital archives, in the form of data, are not.

To conclude, I would like to reiterate that it is important and timely to advance debates on military, humanitarian, legal, and ethical questions posed by new technologies of warfare. I am delighted to see that in this room, we have people with great expertise on cyber technology, artificial intelligence, autonomous weapon systems, outer-space operations, and the use of new technologies in humanitarian operations. We have experts with a technological background, we have military operators, we have policy makers, we have humanitarians, and we have lawyers. I hope that we will learn from each other and be able to advance discussions on IHL implications of new technologies of warfare.

Opening remarks

Sebastiano CARDI
Director General for political affairs and security of the Ministry of Foreign Affairs and International Cooperation of Italy

It is an honour for me to deliver some remarks today on behalf of the Italian Ministry of Foreign Affairs and International Cooperation. I would like to thank the promoters of this event for having once again, this year, gathered such an influential group of experts to discuss issues of exceptional relevance and interest that require further attention, notably the new challenges and developments in the field of International Humanitarian Law (IHL).

The respect of IHL is not only a legal obligation but also a moral imperative: promoting its widest application is crucial in order to ensure that the principles of humanity and dignity of every human being, especially those belonging to vulnerable groups, are always protected even in situations of conflict. Italy continues, therefore, to be at the forefront in the defence and promotion of International Humanitarian Law, particularly to guarantee the highest protection for civilians during today's armed conflicts.

This Round Table meeting is an important opportunity to draw the attention to these matters and to strengthen our commitment for International Humanitarian Law, which is currently facing new and complex challenges, particularly due to the continuous development of technologies.

I find it particularly relevant and timely that this Round Table is focused on the implications posed by the design and deployment of Artificial Intelligence applications and autonomous weapon systems in military operations, as well as on the concerns raised by the growing offensive capacities developed in the field of cyber and space domains. We need indeed to increase our collective understanding of this complex topic of emerging technologies, and their possible military use, in order to avoid a scenario whereby rapid advances outpace our ability to maintain human control on crucial functions of weapon systems, and ultimately our ability to uphold International Humanitarian Law.

I would like to express my profound appreciation for the activities carried out by the International Institute of Humanitarian Law in Sanremo that since its establishment in 1970 has played a crucial role in promoting

International Humanitarian Law, representing a real "*centre of excellence*" at international level. The Ministry of Foreign Affairs and International Cooperation has a longstanding relationship with the Institute of Sanremo. We are committed to supporting and promoting the activities of the Institute, in particular the specific training on international humanitarian law carried out in favour of members of the armed forces of many countries.

I hope you have good and fruitful discussions in the next few days.

Message

Peter MAURER
President, International Committee of the Red Cross (ICRC)

Interviewed by Helen Durham

Helen DURHAM:

Distinguished Guests, I'm very pleased to be here with the ICRC President, Mr Peter Maurer, to talk about the humanitarian implications of new technologies in warfare. New technologies, or digital transformation as we often say, are deeply changing our lives across the sectors and I am personally very fascinated by some of the new innovations that the humanitarian sector, and in particular ICRC, are implementing in space, for example, virtual reality, where we have created all sorts of exciting opportunities to use this new technology, to train and engage on issues such as visiting prisons, training and first aid and forensics. In this sense, what implications do you see, Peter, for new technologies, specifically in the humanitarian sector?

Peter MAURER:

Let me first and foremost give a warm welcome to all the participants of this year's Sanremo Conference. I find it really encouraging that such a topical issue will be given space for debate over the next few days. As you rightly say, Helen, it strikes me too that technological change changes the humanitarian environment and humanitarian work quite fundamentally and this is not any different from any other part of society. While maybe a couple of years ago, when we spoke about technological change and technological transformation people thought: "it's just another computer in your office". You'll realize that this is a fundamental societal change which changes the way we do things. And as you rightly say, we have to unpack a little bit to see what this really means. The edification of every aspect of life has a major impact on humanitarian work. Analytically, big data analysis changes the way we are able to do humanitarian work because we have a much more granular view on how needs are evolving, where needs are and where we have to focus our priorities. Technological change, that edification of our environment, also changes the relationship between us as humanitarian agents and agencies and beneficiaries of humanitarian assistance: Beneficiaries have a much more direct contact amongst

themselves, towards donors, towards the world outside. So, the intermediation of the relationship between humanitarian agencies and humanitarian beneficiaries is happening as we speak. Then, when we deliver humanitarian assistance, and that was your example of virtual reality, it also changes the way we work, the way we deliver. If we didn't have new technologies we would not be able to deliver cash services to people on cellphones; we would not be able to teach international humanitarian law (IHL) by creating virtual reality and by creating another atmosphere for teaching; we would not be able to use data in a much more computing way in order to reunite families. So, there are hundreds of applications now of technological change in humanitarian work, and it goes from law to operations to policy issues which we're dealing with.

Helen DURHAM:

Yes, thank you Peter. As you have said so clearly, new technologies hold great promise, but I think we also need to acknowledge that they also pose a number of risks and I would say in particular perhaps that this seminar will focus on the risks around methods and means of warfare. From your point of view, what are the main issues and perhaps risks and challenges that we need to look at when it comes to new technologies and specifically warfare?

Peter MAURER:

It strikes me when I talk to militaries around the world, at least in some countries, the first thing that comes to their mind is the huge advantage of technological change, as it holds the promise of targeting, of accuracy, of compliance with IHL. While, of course, it raises a lot of issues as to whether this promise is really happening and how enhanced weapons technology is changing the humanitarian landscape in which we are. We know that this is a very ideological debate. It's a substantive debate, it's a polarizing debate because, wherever people come from, whether they are militaries, humanitarians, potential victims, they look at the risk landscape. When they see autonomous weapons, they see the idea that human control is lost in the process of technological change and of the changing environment of warfare. This raises fears and it raises, of course, complicated legal questions, as we know, on how to frame human control that seemingly most of the participants in that debate wish to be maintained in the future. But what does human control really mean? And can we have accurate legal framing and accurate evidence that human control can

somehow bring technological change to minimize risks and to maximize advantages?

Helen DURHAM:

There will be quite a lot of debate on this issue of autonomous weapons in the next few days at this Round Table. In just a few words, what is the ICRC's position currently on autonomous weapon systems?

Peter MAURER:

As I alluded before, I think it is of critical importance to establish what exactly autonomy is and where exactly human control comes in. It raises, of course, complicated legal issues on how to frame it but also ethical issues: if technology holds the promise of enhancing the accuracy of weapons and if elements of autonomy are now being introduced to weapon systems, where is the element of control and how do you frame that? It raises ethical issues: should machines take decisions? And if so, what kind of decisions and who is ultimately responsible in that process? I think this is the crux of the matter and I think the critical issue is indeed to frame human control and to ask the ethical questions. Whether this should happen or not is not the legal question, this is a deeply political question which will be debated in the political environments. People will have opinions, not only with regard to international humanitarian law, but also as to whether machines should take decisions and be programmed in a way they can take decisions autonomously. I think we have seen this, from autonomously driven cars to autonomously driven weapons, that these are very emotional, ethical and legal issues which need to be debated at the same time.

Helen DURHAM:

Thank you. It sounds like we've got a little bit of work ahead of ourselves but on a very interesting topic. Just finally, when we think about the governance structure for, for example, the development of new weapons and technologies, and we have the 70th anniversary of the Geneva Conventions this year, what do you think, going forward, would be needed to make sure that IHL remains relevant in this newly digitized technological age?

Peter MAURER:

I think over the last couple of years I have really advocated that we take a proactive role in interpreting the Geneva Conventions so as to logically establish an adequacy between the Conventions and the reality in which we

are. We have to get to grips with what the terms invented, defined, framed in the Geneva Conventions, mean in cyberspace. I think this work brings us, to a certain degree, to ensure that we build on the past and that we don't discuss artificial gaps which may not exist because, through adjusting the law to reality and by interpreting the law, we can obviously do a lot to clarify what the situation is today. For me there is no question that the Geneva Conventions are and will be highly relevant if technological change comes and is paired with kinetic power in warfare today and that the Conventions are applicable. But then there is this space where we see humanitarian impact outside traditional armed conflicts. This needs to be framed and the intersection of cyberwar and cybersecurity needs to be debated, clarified and also thought through to define which legal system we have to refer to for which kind of situation. At the end of the day, the core issue in these modern warfare and cybersecurity issues is attribution: if something happens and has humanitarian impact where is the origin of it and who is responsible for it? We all know that this is a critical issue which needs further debate, further framing and which is not once and forever written into law. I think the critical issue is to have this debate where we translate existing legal frameworks into new realities and we identify as precisely as we can the gaps and then we will have the complicated question as to who will sit at the table to discuss these issues. My sense is, contrary to 1949, we need to have other participants at the table as well: we will have to have tech companies informing the debate; and we will have to have societies bringing ethical questions to the debate. Therefore, my preview is that international humanitarian law in cyber warfare cannot be debated as a specialist branch of militaries and humanitarians in the future. Other people will have to and will raise their voices and we have to accommodate them and be ready to listen to them and to see what a reasonable development of legal frameworks in that new world is.

Helen DURHAM:
Thank you, Peter, thank you very much.

Peter MAURER:
Thank you.

Keynote address

International Humanitarian Law: Changing and Unchanging 70 Years after the Geneva Conventions

Yoram DINSTEIN
Emeritus Professor, University of Tel Aviv; President of the United Nations Association of Israel; Council Member, IIHL

Change and the Geneva Conventions

On the occasion of the 70th anniversary of the 1949 Geneva Conventions for the Protection of War Victims, I would like to address the theme of the dichotomy between change and lack of change in international humanitarian law with special emphasis on the Geneva texts.

The importance of the topic comes into relief against the background of an ascendant tendency to treat the Geneva Conventions (now entrenched in popular culture) with reverential worship. Some lay persons appear to regard them as impervious to change and engraved in stone, lacking altogether elasticity and pliability. Such a perception of the Geneva Conventions is entirely wrong in terms of both their starting point and their later trajectories.

At the moment of their inception in 1949, the Geneva Conventions already represented change in the pre-existing law. Two of the four Conventions (the Second and the Fourth) were new. As for the other two (the First and the Third), whereas they followed the trail of previous Geneva instruments dated 1929, they too were marked by innovations.

The First Convention on the wounded and sick in land warfare was no less than a fourth rendition of the seminal wording crafted in 1864 (the intermediate revisions done in 1906 and in 1929). While much of the First Convention trod familiar ground, it nevertheless contained starkly new clauses like Common Article 3 to which I shall refer later.

The Second Convention transferred into the "Geneva Law" a Hague Convention - No. X of 1907 – that, in itself, had been an adaptation to maritime warfare of the Geneva Convention of 1906. For its part, the 1907 text was a rewrite of Hague Convention No. III of 1899 constituting an adaptation of the original Geneva text of 1864.

The Third Convention on prisoners of war introduced into the earlier version of 1929 radical amendments deemed indispensable by dint of the dire lessons learned in World War II.

The Fourth Convention for the protection of civilians, currently considered the fulcrum of the Geneva legal regime, was completely novel in 1949. Its *raison d'être* was the unspeakable Nazi atrocities against civilians perpetrated in the course of World War II.

Thus, the common denominator of all four Geneva Conventions was the desire of the Diplomatic Conference of 1949 for significant reform in international humanitarian law.

When one examines the post-1949 era, it is plain to see that the four Geneva Conventions have by no means frozen in time: indeed, they have undergone permutation both formally and informally. Formal changes occurred in 1977, when the Geneva Conventions were supplemented by Additional Protocols I and II governing the conduct of hostilities (in international and non-international armed conflicts, respectively). Informal changes in the thrust of some stipulations of the Conventions were engendered by subsequent practice; and I shall illustrate this in due course.

The moral of the story is that the Geneva Conventions are and have always been attuned to a constant need for reappraisal.

Law and change

Far be it for me to suggest that change is a unique feature of the Geneva Conventions. Life is about change, and so is the life of the law. No legal *status quo* can be maintained perpetually. Law is a living organism and as such it must evolve. When a legal system does not readjust itself in tandem with changing circumstances, a gap will be created between law and reality. In the long run, such a gap will be catastrophic to the law by eroding its bedrock, namely, societal respect for the law.

If this is true of all law, it is particularly true of international humanitarian law. Every war (*bellum*) becomes a crucible for forging new *jus in bello* in light of the experience gained in the battlefield. The introduction of novel methods or means of warfare prompts a fresh look at the law in force.

This is epitomized in the impetus for the revision of the Geneva Conventions in 1949. In their new configuration, the Conventions reflected an immense pressure brought to bear by public opinion upon governments

to reform a legal system that had been weighed in the balance and found wanting during World War II.

That said, we must be cognizant of the fact that change in the law cannot be unlimited and out of control.

How much change?

The pivotal question is: how much change do we wish for? The quandary relates to the degree of change that can be absorbed by international humanitarian law without a melt-down. *Au fond*, when governments are urged to revise binding norms, it is incumbent on them to assess the rectification sought in a manner congruent with the unchanging (i) axiom; (ii) objective; and (iii) cardinal principles of the legal system.

In case you think that what I am saying now is too abstract, let me give you a concrete example of the ramifications of failing to keep a lid on change. Some air force enthusiasts are clamoring for a new rule of warfare allowing the bombing of enemy civilians, with a view to shattering their morale and bringing the war to a rapid end (thereby, in the final analysis, perhaps saving incalculable numbers of lives of both civilians and combatants). Factually, not much evidence can be adduced for the proposition that – by themselves - "shock and awe" attacks against the civilian population will force a government to capitulate. In World War II, devastating Allied air strikes in Germany and in the Pacific – pulverizing and even incinerating whole cities – failed to achieve the purpose of the architects of "strategic bombings". Still, let us assume *arguendo* that large-scale bombings of the civilian population might terminate a war promptly. Could they be reconciled with the existing *jus in bello*? The answer is that direct attacks against civilians – even if effective in practice – are patently incompatible with the cardinal principle of distinction that I shall dwell upon. Given the cardinal nature of the principle, recasting international humanitarian law along the lines proposed would be inherently impermissible.

Unchanging axiom of international humanitarian law

The first obstacle to any change in international humanitarian law is the axiomatic major premise of the equal applicability of the *jus in bello* to all Belligerent Parties in an international armed conflict, irrespective of who is

the aggressor (and, correspondingly, who is the victim of aggression) under the *jus ad bellum*. This axiom, like all axioms, is a basic postulate with which no rule can be in disharmony.

In many Universities in the world, Professors of Ethics teach the so-called "just war theory", whereby if a war is "unjust" that taints the status of combatants in the conduct of warfare. Since Ethics is not my discipline, I shall not encroach into it. But, from the standpoint of international law, the linkage between the legality of war and the conduct of hostilities is utterly unacceptable. Whether the war is "just" or "unjust" under the *jus ad bellum*, all combatants are equally bound by the same obligations and enjoy the same rights pursuant to the *jus in bello*.

Think about it against the backdrop of World War II. Despite the fact that the Nazi war of aggression was singularly unjust and unlawful, German combatants who fell into the hands of the Allies on the Western Front were still entitled to the privileges conferred on prisoners of war, in conformity with the Geneva Convention of 1929 (which, regrettably, was not applicable on the Eastern Front). Were it not for the profound disconnection between the *jus in bello* and the *jus ad bellum*, millions of German captives could have been excluded from the benefits of international humanitarian law.

The axiom of equality of the Belligerent Parties is viable today even when UN forces are engaged in an armed conflict: the same *corpus* of international humanitarian law binds these forces (representing the international community) and their opponents (whoever these opponents are).

Interestingly enough, the axiom of legal equality of the parties is apposite also to non-international armed conflicts, although there is no *jus ad bellum* regulating such conflicts, and besides there is a built-in disparity in the positions of the two principal adversaries (the government and insurgents organized armed groups). Equality of the parties denotes that insurgents – no less than the government - must apply international humanitarian law, including the prohibition imposed by Common Article 3 of the Geneva Conventions on sentencing accused personnel "without previous judgment pronounced by a regularly constituted court, affording all the judicial guarantees". Insurgent armed groups must abide by this norm whether or not they be in effective control of the territory. Of course, in the absence of control of any territory, insurgents cannot conceivably ensure the operation of a regularly constituted court affording all the judicial guarantees. Still, they are not granted a dispensation from the rule: they will consequently be barred from sentencing offenders.

Unchanging object and purpose of international humanitarian law

The object and purpose of the *jus in bello* were lucidly proclaimed already in the St. Petersburg Declaration of 1868 (four years subsequent to the adoption of the original Geneva Convention): "alleviating as much as possible the calamities of war".

Unfortunately, war is not a game of chess: it always entails the spilling of blood and the destruction of property. What international humanitarian law does is balance – "as much as possible" - military necessity with humanitarian considerations. The outcome is a compromise between these polar opposites. On the one hand, war is pursued with the goal of winning it; on the other, it is of the essence of the *jus in bello* that not everything is allowed in war.

A more recent formulation of the same fundamental idea – contemplated from a different angle - is enshrined in Article 35(1) of Additional Protocol I of 1977, prescribing that "the right of the Parties to the conflict to choose methods or means of warfare is not unlimited". Indubitably, the primary precept promulgated in Article 35(1) accurately reflects customary international law (which is binding also on non-Contracting Parties to Additional Protocol I). Wartime cannot become a "kill-free" temporal domain.

Unchanging cardinal principles of international humanitarian law

As the International Court of Justice famously pronounced, in the 1996 Advisory Opinion on the *Legality of Nuclear Weapons*, international humanitarian law is underpinned by two cardinal principles:

(i) Distinction (between combatants/military objectives and civilians/civilian objects); and
(ii) Avoidance of unnecessary suffering or superfluous injury (to combatants).

It must be fully appreciated that these two cardinal principles are like legs on which international humanitarian law is standing. Were they to be amputated, the whole body of that law would collapse.

The cardinal principle of distinction is conspicuously momentous. I have already adverted to it in the context of the insupportable notion of "shock and awe" air strikes directed against civilians with a view to shattering morale. Let me add that the principle of distinction protects the

civilian population not merely from deliberate attacks (whatever their underlying rationale) but also from indiscriminate attacks that are oblivious to the identity of the potential victims (be they combatants or civilians). That leads me to the principle of proportionality.

The principle of proportionality

The principle of proportionality is derived from the cardinal principle of distinction and is an extension of the prohibition of indiscriminate attacks. In the actuality of modern warfare – as waged, at least, by the armed forces of the countries represented in this auditorium – the principle of proportionality has become the gravamen of the protection of the civilian population from injury.

The dilemma of proportionality comes into play when lawful targets (combatants/military objectives) are attacked, instigating collateral damage (or incidental loss) to civilians/civilian objects. The trouble is that civilians/civilian objects cannot be comprehensively insulated from any form of collateral damage or injury, inasmuch as they are almost always present in or near combatants/military objectives. The only exceptions would be attacks mounted in the middle of the desert, in mid-ocean or on the arctic ice-cap, and even then civilians may turn up in the vicinity by chance.

International humanitarian law takes this stubborn fact of life into account. What it lays down – in the form of the principle of proportionality – is that an attack against a lawful target is proscribed if it is expected to cause collateral damage to civilians/civilian objects, which would be "excessive" in relation to the concrete and direct overall military advantage anticipated.

The main feature of the principle of proportionality is the expectation of "excessive" collateral damage to civilians/civilian objects. This is a matter of foresight rather than hindsight: what counts is what is expected in advance of the action on the footing of a reasonable evaluation of the information available at the time. Moreover, "excessive" does not mean "extensive": lawful collateral damage to civilians/civilian objects may be quite extensive, if – but only if – it is commensurate with the anticipated overall military advantage.

A cautionary note: the principle of proportionality must not be read beyond the ambit of the protection of civilians/civilian objects from

"excessive" collateral damage. No proportionality is required as regards death, injury or destruction inflicted on combatants/military objectives.

What does – and must – change?

The crucial question, therefore, is: what changes in international humanitarian law are appropriate – indeed, inevitable - and what changes are inadmissible? The axiom remains immutable. The object and purpose are enduring. The cardinal principles are firmly fixed. However, while all these are unswerving, there is an abundant subsidiary body of international humanitarian law that is altered incessantly.

Each generation reinterprets the same cardinal principles differently, coming up with new solutions to both old and new problems. The principle of proportionality – playing such an important role today – is emblematic in having provided a new solution to the old problem of collateral damage. During World War II, it was still possible to rationalize an attack causing devastating carnage to enemy civilians/civilian objects – highlighted by the atomic bombing of Hiroshima – on the ground that that the city constituted a military objective. This was the case because Hiroshima was an important seaport, serving as a supply center for the Imperial Japanese military, with several thousand troops stationed there. A classification of a target as a military objective seemed at the time to close the book on the legal analysis of a projected attack. But the principle of proportionality, as developed in the post-War era, recalibrates that analysis. The "excessive" collateral damage to civilians (through blast, heat and radiation) would render illegal today a Hiroshima-like attack, even though it was directed at a military objective.

New *jus in bello* problems arise as a result of either technological developments or a shift in battlefield tactics, and they may make it vital to review and update obsolete rules. There is nothing wrong with such reviews and updates – indeed, these may be ineluctable - as long as they do not tamper with the essential components of international humanitarian law.

How is change brought about?

When there are compelling reasons for change in the rules of international humanitarian law, what is the process by which the change

can be generated? In this respect, the *jus in bello* is no different from all other branches of international law: change is spawned either by treaty or by custom.

Custom means the general practice of States (with an accent put on "specially affected" States) plus *opinio juris*. Custom is unwritten, and it may not be easy to pinpoint. By contrast, a treaty is a written agreement between States. A treaty has an advantage over custom in being *jus scriptum*, but – unlike custom, which is commonly binding on the entire international community – a treaty is binding only on Contracting Parties. Here is where the Geneva Conventions stand out, since they have been ratified or adhered to by every country in the world. A handful of other treaties (primarily, the Charter of the United Nations) are on the cusp of universal acceptance; but, so far, only the Geneva Conventions have achieved that goal.

A treaty may be innovative, deviating from pre-existing customary law in the relations between Contracting Parties (without affecting third States). Conversely, a treaty - in whole or in part - may be, or may become in time, declaratory of customary international law. In that case, the declaratory norms (by virtue of their customary nature rather than owing to the treaty) are binding also on non-Contracting Parties. In the sphere of international humanitarian law, this is a dominant issue whenever edicts of Additional Protocol I of 1977 are relied upon. The perennial question is whether a relevant clause of the Protocol is accepted as declaratory of customary international law. If it does not reflect custom, the provision will usually be contested by non-Contracting Parties (led by the US).

Change in international humanitarian law can be attained not only through treaty, but also by means of customary evolution. Irrefutably, a new custom may modify a previous custom. Furthermore, a new custom may make inroads into a treaty text by reinterpreting it in keeping with subsequent practice. While - on the face of it - the text remains intact, the substance of the norm will undergo a significant transformation in reality. I shall illustrate this, in the context of the Geneva Conventions, in a moment.

How is change *not* brought about?

It is of salient importance to underline that change in international humanitarian law can be brought about solely by States acting jointly under the banner of a treaty or custom. There is no other way to validly effect such change.

This should put in proper perspective the role of the "civil society" in the process. By themselves, non-governmental organizations (including even the foremost non-governmental organization, *viz*. the ICRC, which is endowed with a special standing under the Geneva Conventions) are incapable of producing change in international humanitarian law. Non-governmental organizations - and other non-State actors – can definitely be instrumental in demanding that governments conclude innovative treaties or reset their practice. The history of the adoption of the 1997 Ottawa Convention on Anti-Personnel Mines shows that - through incessant goading of governments - non-governmental organizations can leave an indelible mark on the unfolding of a groundswell of State support for the creation of new humanitarian norms. But not always do projects championed by the "civil society" come to fruition. Even when they do, the part played by non-governmental organizations is strictly a behind-the-scenes performance. Success in promoting a treaty does not turn non-State lobbyists into accredited members of the cast of actors on the law-making international stage.

A single State is equally unable to convert international humanitarian law unilaterally: it takes a group of States to produce change collectively. If an individual State acts in breach of either a custom or a treaty by which it is bound, the breach remains a breach as long as the acting State is not joined by other States in defying the law. Contrarily, once a number of States share a policy of disapprobation of a law in force, the legal landscape is liable to transmute. At the end of the day, it may be conceded that the first breach of the law was merely a building-block of what has ultimately turned into subsequent practice reshaping a pre-existing treaty or custom. So, the issue of change versus lack of change of a specific norm is not as simple as it sounds at first blush. There are occasions when one has to take a pause and perhaps wait a few years before it is known conclusively whether an incipient breach of the law has (or has not) ripened into an unstoppable subsequent practice.

I shall give two illustrations of subsequent practice impinging on ostensibly sacrosanct provisions of the Geneva Conventions: one relates to the Second Convention and the other to the Third Convention.

Subsequent practice – second Geneva Convention

Under Article 34(2) of the Second Geneva Convention, "hospital ships may not possess or use a secret code for their wireless or other means of

communication". The wording can hardly be clearer, yet it no longer represents the law prevailing at the present time. Even the new (2017) ICRC commentary on the Second Convention admits that, bearing in mind State practice since 1949, the mere possession or use of secret codes aboard hospital ships cannot anymore be denounced as a breach of international humanitarian law (unless the use is harmful to the enemy).

Credit ought to be given where credit is due. The practical problem with Article 34(2) was first identified in Rule 171 of the San Remo Manual of Maritime Warfare (finalized, under the auspices of this Institute, in 1994). By the time of the drafting of the San Remo Manual, it became apparent that failure to receive encrypted communications would jeopardize the ability of hospital ships to function effectively. Since then, the gap between the law and reality has further widened, for naval code messages have totally replaced communications *en clair*. The San Remo Rule was still hesitant in its language, given natural qualms about challenging a Geneva text. A quarter of a century later, the hesitation is no longer cogent. It is indisputable that the only restriction today on the use of cryptographic equipment on board hospital ships is the prohibition of abuse, and the sole need is to ensure the non-transmission of coded messages harmful to the enemy (e.g., intelligence data).

Subsequent practice – Third Geneva Convention

Article 118 of the Third Geneva Convention decrees that prisoners of war "shall be released and repatriated without delay after the cessation of active hostilities". The nub of the matter is the phrase "and repatriated", which was adopted in 1949 after ample consideration. Indeed, an Austrian-sponsored amendment - granting released prisoners of war the option not to return home if they did not wish to do so - was rejected at the time.

This became a major bone of contention in the Korean armistice negotiations (it is noteworthy that the Korean War broke out less than a year after the crafting of the Third Convention). The controversy arose as a result of a massive refusal of North Korean and Chinese prisoners of war to return home. Hostilities went on purposelessly for two years – precipitating many casualties with trivial gains for either side on the ground – until, in 1953, the Parties to the conflict agreed not to coerce released prisoners of war to be repatriated involuntarily (an intricate scheme was worked out to verify the true wishes of those released from imprisonment).

The Korean precedent was followed in the Gulf War in 1991, and the exclusion of compulsory repatriation of prisoners of war (without prejudice to any agreed-upon formula delineating the mode of their release) may now be viewed as binding customary law. Thus, Article 118 of the Third Convention should no longer be taken as read.

Restatements and change

Since 1977 – the date of the adoption of Additional Protocol I, which has left in its wake a great deal of bitter disputes – States have been overtly reluctant to indulge in new treaty-making efforts in the field of international humanitarian law, except where certain weapons are concerned. Whereas means of warfare are the gist of sundry post-1977 treaties, there has been no new treaty germane to methods of warfare**.** Every once in a while, appeals are made for an innovative treaty addressing this or that manifestation of the *jus in bello*, but to date none of the initiatives has been crowned with success. The bottom line is that it is not likely that any novel treaty on methods of warfare will emerge in the foreseeable future.

What is the alternative? The emphasis has shifted from treaties to custom. And, in order to articulate customary law in an authoritative up-to-date manner, a technique has evolved of preparing non-binding restatements of the law in the form of manuals. Such restatements/manuals are the products of groups of experts consisting of both academics and practitioners, collaborating in their individual capacity albeit in some consultation with governments of core States (as well as the ICRC). The prototype of restatements/manuals was first moulded in the 1994 San Remo Manual on Armed Conflicts at Sea. Since then, we have had several additional restatements/manuals on selected problems of international humanitarian law – for instance, Air and Missile Warfare – all emulating the San Remo model.

A critical dimension of restatements/manuals is that, to be useful, they must predominantly mirror the *lex lata*. After all, unlike official organs of States who may devise new law, experts are not qualified to do so. There is no genuine value added in a restatement/manual reflecting the *lex ferenda* from the experts' standpoint. What experts wish for may be interesting as a moot academic exercise, but it is of little empirical use to the end-users of their product (military operators and their legal advisers). What the experts have to do is examine the actual practice of States, trace patterns of behaviour, and conclude by portraying the law as it is. The experts may

illuminate a burgeoning trend in the practice of States that may be indicative of a law in the offing (*in statu nascendi*). Still, the experts' paramount task is to establish what the law is *de lege lata*.

The search for *lex lata* means identifying any applicable custom; invoking all relevant treaties (expounding their innovative or declaratory status); explaining when custom (obligatory for all States) and treaties (binding as such only on Contracting Parties) are at odds with each other; construing ambiguous treaty provisions; and trying to elucidate diverse diagnoses of the general practice of States.

A restatement/manual must be *au courant*: it has to show when an article or a paragraph in a treaty (which was innovative at the outset) has generated new custom, and (the other way around) when subsequent practice has modified a treaty clause. This is a complicated mission, and not always do the experts manage to build a consensus. Unresolved disagreements must be reflected in a commentary accompanying the black-letter rules of the restatement/manual. The commentary will also shed light on the choice of words in drawing up those rules and include cite-references to treaty texts underlying them.

Naturally, no less than other texts, a restatement/manual must be reconsidered after a reasonable lapse of time. A restatement/manual, even if flawless when inaugurated, is liable to lose its cutting edge over the years. It therefore has to be periodically reviewed through the lens of any posterior growth of international humanitarian law. Accordingly, the San Remo Manual - after a quarter of a century of successful existence (in the course of which it has been cited countless times) - will soon be undergoing re-examination by a new group of experts. The reason is plain to see: whether or not the San Remo Manual was 100% perfect in 1994, it can scarcely be a 100% perfect a quarter of a century later. Subsequent practice must be reckoned with.

The pace of change

The pace of change in customary international law is usually slow, yet the rate may accelerate very swiftly. Although the phrase "instant custom" is an oxymoron, sometimes custom can consolidate over a relatively short period of a few years. The remarkable development of the customary law of the sovereign rights of a coastal State in its continental shelf, within a single decade from the debut of this construct, is a paradigmatic example. And one of the best illustrations of a quickening momentum of customary

germination can be elicited from international humanitarian law itself where non-international armed conflicts (NIACs) are concerned.

For a long stretch of time, there was simply no international humanitarian law governing NIACs. The genesis of NIAC humanitarian law is to be found in 1949, in Common Article 3 of the Geneva Conventions. This was a very pregnant moment, but it must be seen in proportion: there was one single clause on NIACs - common to all four Conventions - compared to more than four hundred other provisions focused on international armed conflicts.

In 1977, Additional Protocol II (devoted exclusively to NIACs) was signed. Additional Protocol II was much shorter and less impressive than its twin, Additional Protocol I (regulating international armed conflicts). But, even in its truncated form, Additional Protocol II had to overcome strong opposition by numerous States.

For more than a decade and a half, it looked as if progress stopped in its tracks, and Additional Protocol II seemed to be destined to remain the final word on NIACs. Then, in the mid-1990s, there was a quantum leap catapulted by the Statutes of the *ad hoc* International Criminal Tribunals for the former Yugoslavia (ICTY) and Rwanda (ICTR), and the spate of Judgments delivered by these Tribunals. Alongside a vigorous move forward of NIAC humanitarian law, individual criminal accountability was attached to serious violations of the law. For the first time, legal breaches were recognized as war crimes when committed in a NIAC setting.

The trend culminated in the 1998 Rome Statute of the International Criminal Court (ICC), which in Article 8 recites a long roster of NIAC war crimes (their listing being augmented further in Kampala in 2010). Thus, there has been a sea-change in the NIAC law. Commencing with zero NIAC humanitarian law prior to 1949 and going through a phase of moderate expansion in 1977, NIAC law has now grown exponentially making headway in the province of war crimes. One can only marvel at a metamorphosis in the legal canvas within a relatively short time span.

Changes in international humanitarian law due to technological developments

Technological developments affecting the means and methods of warfare inexorably require constant changes in international humanitarian law. That has always been the case. The invention of war planes and missiles, the introduction to land warfare of tanks, etc., all contrived to alter

the *jus in bello*. There was no way for the law to ignore the repercussions of the new destructive capabilities of Belligerent Parties. Currently, we are going through a period in which the march of endless technological developments is even more pronounced: the tempo gets faster and faster from one generation to another.

Prime illustrations of current technological developments are: (i) cyber warfare; (ii) semi-automated weapon systems; (iii) drones (remotely piloted aircraft); (iv) unmanned maritime surface vessels and underwater devices; (v) land robots; and (vi) satellites in outer space. These will be discussed in detail at the present Round Table.

There are a host of lawyers who desire to scrutinize also the impact on international humanitarian law exerted by artificial intelligence, a technology that still has quite a distance to go before it fully materializes. For my part, I find it premature to thrash out, e.g., the penal consequences of a robot determining by itself whether collateral damage to civilians/civilian objects is expected to be "excessive" compared to an anticipated military advantage of a lawful attack. The issue of human accountability for rogue actions by futuristic autonomous contraptions is undeniably fascinating. All the same, I would prefer to let technology advance further before its legal reverberations are submitted to a coherent legal discourse.

Technological developments improving the ability to comply with international humanitarian law

Technological developments are usually looked upon as impediments to the implementation of the existing *jus in bello*, and (as a corollary) catalysts for relentless change. But it must be perceived that exceptionally technological developments may also lay the ground for a better implementation of the law in force, thus fending off any incentive for change.

Two leading examples should suffice. The first is the use of PGM enabling an attack against a lawful target to have a surgical effect, thereby alleviating the danger of indiscriminate bloodshed and substantially minimizing – possibly eliminating altogether - collateral damage to civilians/civilian objects.

Secondly, surveillance drones (remotely piloted aircraft) can furnish real-time information about the presence of civilians/civilian objects in proximity to a military objective. People are disposed to think of drones as

weapon-carrying aerial platforms. Yet, by far the large majority of drones in use today are surveillance drones. By loitering over a prospective target, drones can overcome the "fog of war" and collate accurate data about the contiguity of civilians/civilian objects. Upon inspecting the intel gathered by the drone, the planners of an attack may reliably determine whether the expected collateral damage would impel aborting the attack. Evidently, there is also the possibility of opting for alternative courses of action, such as embarking on the attack in a different timeframe (say, night-time) or in a divergent mode.

Use of latent technological developments

Not all technological developments have direct palpable effects, for better or worse, on the *jus in bello*. There is a raft of technological developments that may influence that law latently or tangentially, depending on context.

Thus, international humanitarian law obligates an attacker to use feasible precautions, including - where possible - the issuance of warnings to civilians about an impending attack. Obviously, warnings to civilians cannot always be released in practice, for surprise may be of the essence of the plan of attack. But, if warnings are feasible, they may serve as decisive precautionary measures, precluding "excessive" collateral damage to civilians.

When feasible, how are warnings to civilians to be issued? In the past, the technology was limited in its range to the use of megaphones; recourse to Radio/TV broadcasts; dropping of leaflets from the air, and so forth. In the electronic age, warnings can also be issued to the civilian population through messages sent by SMS, via the "social networks" (such as Facebook) on the Internet, etc. Large numbers of civilians who were once literally beyond reach can now be effectually contacted. Thus, new technologies designed for normal peacetime purposes may induce unforeseen benefits in wartime by safeguarding civilians from some collateral damage.

Changes unrelated to new technologies

New technologies are not the only roots of change in warfare or in the *jus in bello*. Innovations are often made necessary by exposure to ever-

changing tactics that are not necessarily linked to any state-of-the-art technologies. Curiously enough, international humanitarian law is apt to find it harder to tailor itself to non-technological challenges.

A rudimentary illustration relates to the use of "human shields" (*i.e.* civilians) to screen lawful targets from attack, a forbidden tactics which of late has become flagrant in multiple armed conflicts. The issue of unlawfully emplacing (voluntary or involuntary) "human shields" in front of or amid combatants/military objectives was probed by a special session in Rome organized by this Institute with Judge Pocar in the Chair. Regrettably, no unanimity has emerged about pragmatic sanctions against such tactics. All that could be agreed upon across the board was that international humanitarian law must somehow come to grips with this test of its authority.

A related issue, not sufficiently studied in my opinion, is the lack of adequate response by international humanitarian law to the mushrooming phenomenon of the use of "suicide bombers" (masquerading as civilians). "Suicide bombers" are launched today all over the world, but international humanitarian law is still baffled by the question of how to deter them. After all, by definition, a "suicide bomber" is bent on suicide; so, it is impossible to deter that person by simply threatening him/her with death. Other means of deterrence are elusive by reason of the existing prohibition of collective punishments against innocent kith or kin.

These and other conundrums posed by deceitful methods of warfare must command proper attention by lawyers and States. An ostrich-like policy of burying our heads in the sand will merely encourage military operators to improvise countermeasures that lawyers may not be happy about. Remember that, however unpalatable, such countermeasures may ultimately insinuate themselves on international humanitarian law under the mantle of subsequent practice.

A new Matrix?

A failure by international humanitarian law to grapple with new challenges ("the decision not to decide") leads to persistent calls in the legal literature for a new matrix. The contention is that, since contemporary conditions of warfare (especially in so-called asymmetrical warfare) are not dealt with in a suitable fashion, there is no escape from the necessity of reshaping the very matrix of international humanitarian law.

As I have argued in this presentation, there is always room for law reform. But I do not believe that the present matrix of international humanitarian law is irremediably defective only because some issues (like "human shields" and "suicide bombers") remain for the time being unsettled. As I see it, even if we are dissatisfied with lack of progress in certain directions, there is nothing that is drastically wrong with the nucleus of international humanitarian law.

In any event, it must be underscored that States – the ultimate stakeholders here - do not reveal any inclination to revisit the fundamental structure of international humanitarian law on account of peripheral deficiencies.

Conclusion

There are three points that I want to bring to the fore in conclusion:

i. International humanitarian law (like all law) invites constant review and updating, and it must be continuously scanned in order to verify whether change has actually occurred (perhaps unnoticed) or is forthcoming. Nevertheless, there is no real need for a revision of its basic tenets.

ii. Since no new general international humanitarian law in the form of a treaty is envisioned any time soon, the challenge of change in this field can only be confronted through evolution in customary international law. Progress by custom can be attested to by restatements/manuals elaborated by experts.

iii. Although change in customary international law is ordinarily a slow process, the NIAC example amply demonstrates that – when the international community is ready and willing – law reform can be speedily accomplished. Quick transformation of international humanitarian law has happened in the past, and it can safely be prognosticated that this will happen again – as and when warranted - in the future.

I. The Geneva Conventions on their 70th anniversary: IHL and the changing realities in the conduct of hostilities in the past century

From International to Non-International Armed Conflicts: IHL and the changing realities in the nature of armed conflicts

Gabriella VENTURINI
President, Italian Branch, International Law Association; Member, IIHL

Introduction

In the second half of last century as well as in recent decades the established bipartition between international armed conflict (IAC) and non-international armed conflict (NIAC) and the respective laws has been seriously challenged. While armed confrontations between states have continued to occur, either individually or through coalitions of states and occasionally with the participation of an international organization (IO), the vast majority of the armed conflicts after 1949 were NIACs and they still outnumber IACs at the present time. Contemporary NIACs, however, do not always correspond to the traditional type of internal conflict – insurgents fighting against the government of a state within the boundaries of its territory – which before 1949, and irrespective of its motives – change of government, secession or others – was considered as falling within the purview of individual states only.

As regards the parties to the conflict, internal armed conflicts may occur either between state armed forces and non-state armed groups (NSAGs) or among different NSAGs. As to their geographical dimension, a number of armed conflicts which are not purely inter-state are not just internal either, because they do not take place in the territory of one single state – and for this reason they are called cross-border, or in some cases transnational, armed conflicts. Sadly, the expansion of armed violence has given way to criminal acts and has caused immense suffering to the civilian population in the affected areas.

From 1949 onwards international humanitarian law (IHL) has developed trying to respond to the changing realities in the nature of armed conflicts on the basis and in the framework of the Geneva Conventions (GCs).

NIACs at Geneva

According to Common Article 2 the GCs apply to IAC, i.e. "to all cases of declared war or of any other armed conflict which may arise between

two or more of the High Contracting Parties, even if the state of war is not recognized by one of them." Yet, while negotiating the GCs the delegates at the Diplomatic Conference were well aware of the NIAC issue. After all, the Spanish Civil War had preceded World War Two by a few years and the past decades had witnessed long and bloody civil wars, like for instance the Mexican Revolution (1910-1920) or the Russian Civil War (1917-1922) as well as a multitude of rebellions and uprisings in different parts of the world.

Although NIACs were deemed to be an intra-state matter, international practice had developed some measures, such as recognition of belligerency or recognition of insurgency, which would produce certain effects on the relations between the parties to the conflict and between them and third states.[1] Moreover, Red Cross Societies and the ICRC had a long record of initiatives and studies aimed at extending IHL obligations, or at least the basic ones, to situations of NIAC.[2]

Various proposals, which would make international humanitarian rules regarding the wounded, sick, shipwrecked, and prisoners of war applicable in NIAC, were discussed at the 1946 Preliminary Conference, at the 1947 Conference of Government Experts, at the Stockholm Conference of the Red Cross in 1948 and eventually at the Diplomatic Conference in 1949.[3] A Special Committee, then a Working Party were established by the Conference and they submitted several drafted versions of the provision which after complex negotiations became Article 3 common to the four GCs of 1949 (CA 3) and which changed the scope of IHL forever.[4]

CA 3 has proved to be a true bastion of IHL, on the basis of which customary law of NIAC has developed and is presently universally recognized. The "mini-Convention" condenses the essential rules of the GCs making them binding on each Party to an "armed conflict not of an international character occurring in the territory of one of the High Contracting Parties" but without affecting the legal status of the parties to that conflict. This is probably the most relevant provision of CA 3, which affirms and enshrines parity in obligations, while acknowledging asymmetry in status.[5]

---

[1] Such belligerent practice is reviewed and analysed by L. Moir *The Law of Internal Armed Conflict*, Cambridge University Press, Cambridge, 2002, pp. 4-18.

[2] The Red Cross initiatives addressing humanitarian concerns in NIACs are recalled in the ICRC Commentary on the First Geneva Convention, 2016, Article 3: Conflicts of not an international character, para.s 362-364. Available at https://ihl-databases.icrc.org/.

[3] *Op. cit.* para,s 366-374.

[4] *Op. cit.* para.s 376-383.

[5] See S. Sivakumaran, *The Law of Non-International Armed Conflict*, Oxford University Press, Oxford, 2012, pp. 242-246. The "revolutionary import at the time" of this provision

CA 3, however, could not govern all different types of situations in a NIAC. The expression "in the territory of one of the High Contracting Parties" would seem to exclude from its scope those NIACs which take place in, or spill over to the territory of different states. A definition of what is an "armed conflict" is not provided. But above all, while protecting "persons taking no active part in the hostilities", CA 3 does not regulate the conduct of hostilities in NIAC. On these and other aspects IHL has developed along two main lines, one of which consists of a normative/legal process and the other one of an interpretive/analytic activity. Faced with the complexities of contemporary NIACs, IHL owes much to two other branches of International Law, i.e. International Human Rights Law (IHRL) and International Criminal Law (ICL) which have contributed substantially to both its normative and interpretive progress.

The normative/legal process

About three decades after the conclusion of the GCs, states undertook an exercise of modernization and strengthening of IHL resulting in the adoption of the two Additional Protocols (APs) of 8 June 1977. With regard to NIACs the outcome of this process was twofold. On the one hand, in the wake of the decolonization period, the "armed conflicts in which peoples are fighting against colonial domination and alien occupation and against racist regimes in the exercise of their right of self-determination" were made subject to the law of IAC.[6] On the other hand, a rather high threshold was established for NIACs other than wars of national liberation, which under AP II were made subject to the basic rules on the conduct of hostilities, including protection of civilian population and of selected civilian objects. Indeed AP II applies only to high intensity NIACs, i.e. armed conflicts not covered by Article 1 AP I and which "take place in the

---

was pointed out by the Constitutional Court of Colombia in its 1995 ruling on the constitutional conformity of AP II at para. 14 (see https://casebook.icrc.org/case-study/colombia-constitutional-conformity-protocol-ii).

[6] AP I Article 1.4. See Y. Sandoz et al. (eds), *Commentary on the Additional Protocols of 8 June 1977*, ICRC, Geneva, 1987, para.s 66–118. It should be noted that at the time of adoption of the APs most wars of national liberation had come to an end. Presently the two main examples are Western Sahara and Palestine. Palestine acceded to the GCs and AP I in April 2014 and in June 2015 the Polisario Front made a unilateral declaration undertaking to apply the GCs and AP I to its conflict with Morocco. While Morocco acceded to AP I in 2011, Israel is not yet a party to the Protocol.

territory of a High Contracting Party between its armed forces and dissident armed forces or other organized armed groups which, under responsible command, exercise such control over a part of its territory as to enable them to carry out sustained and concerted military operations and to implement this Protocol." [7] AP II, however, does not replace CA 3, which continues to apply to the broader range of NIACs going beyond internal disturbances and tensions.[8]

Both APs have incorporated principles belonging to the realm of IHRL. For example, Article 75 of AP I lists a number of provisions which are contained in HRL instruments; while, however, human rights treaties include clauses permitting derogation in times of war, no derogation or suspension of guarantees established in Article 75 are allowed.[9] The influence of IHRL is especially relevant in respect of AP II, the preamble of which expressly recalls that "the international instruments relating to human rights offer a basic protection to the human person." For example, Article 6 AP II applying to penal prosecutions was clearly inspired by Articles 14 and 15 of the International Covenant on Civil and Political Rights (ICCPR) of 1966.[10] Thus, IHRL fertilizes IHL through normative lending.

It should also be added that AP II does not exhaust the normative/legal process of development of IHL in the face of the increase of contemporary NIACs. Indeed, after the entry into force of the GCs several treaties on the protection of cultural property as well as conventions regarding weapons or prohibiting child soldiers have extended their scope of application to NIAC.[11] This is how a gradual harmonization of the law on the conduct of hostilities in IAC and in NIAC has begun.

---

[7] AP II Article 1.1.

[8] See Y. Dinstein *Non-International Armed Conflicts and International Law*, Cambridge University Press, Cambridge, 2014 p. 8.

[9] See Sandoz et al. *Commentary on the Additional Protocols* supra n. 6 at para. 3092.

[10] See Dinstein *Non-International Armed Conflicts* supra n. 8 p. 143.

[11] See the 1954 Convention on the protection of cultural property in the event of armed conflict (CPCP) at Article 19.1 and its 1999 Second Protocol at Article 22.1; the Convention on certain conventional weapons (CCCW) as amended in 2001 at Article 1.2 and its Protocol II on prohibition or restrictions on landmines, booby-traps and other devices as amended in 1996 at Article 1.3; the 1993 Chemical Weapons Convention (CWC) at Article I.1; the 1997 Ottawa Convention on landmines at Article 1; and the 2000 Optional protocol to the Convention on the rights of the child at Article 4.

The interpretive / analytic activity

The development of IHL in the decades subsequent to the Geneva Conventions, and particularly after the adoption of the two APs, benefited greatly from the judicial activity of international and (to a lesser extent) domestic tribunals.

The legal qualification of an armed conflict is almost always controversial, especially with regard to internal conflicts which, in relation to the applicable IHL instruments, may be classified into three different categories: CA 3 NIACs, AP II NIACs and AP I wars of national liberation. A significant example is offered by the first conflict in the Chechen Republic of the Russian Federation (1994-1996). While a rare domestic judicial decision of the Russian Constitutional Court stated on 31st July 1995 that Protocol II was one of the sources of law relevant to the conflict, the Chechen side has for a long time claimed that the war was a conflict governed by article 1.4 AP I.[12] As a matter of fact, the majority of contemporary internal conflicts are deemed to fall under the scope of CA 3.

While neither the GCs nor the APs include grave breaches of CA 3 or violations of other rules applicable to NIAC, in 1993 the Statute of the International Tribunal for the former Yugoslavia (ICTY) first extended individual criminal responsibility to violations of CA 3 and other customary rules in NIAC. From then on a common thread has run through Article 3 of the ICTY Statute, Article 4 of the Statute of the International Tribunal for Rwanda (ICTR), corresponding articles in the statutes of mixed tribunals and eventually Article 8.2(c) and 8.2(e) of the Statute of the International Criminal Court (ICC). It resulted in the flourishing of interpretive activities, which have complemented the IHL applicable to both IAC and NIAC while promoting its knowledge and dissemination among legal scholars as well as the general public.

International criminal tribunals have not been reluctant to address issues of qualification of armed conflicts and interpretation of the applicable IHL. In the judgments and decisions of the ICTY, the ICTR, the mixed tribunals and eventually the ICC we find a comprehensive definition of what

[12] See P. Gaeta 'The Armed Conflict in Chechnya before the Russian Constitutional Court' EJIL Vol. 7, 1996, pp. 563-570 at pp. 568-569. An amendment to the Russian federal act on damages for soldiers deployed in missions to extremely dangerous areas, passed on 19th December 1997, also made reference to the non-international conflict in the Chechen Republic. See M. Mísová 'The legal character of the conflict in Chechnya' 8th May 2001, available at: https://reliefweb.int/report/russian-federation/legal-character-conflict-chechnya.

constitutes an armed conflict,[13] as well as the application to hostilities in both IAC and NIAC of the principles of distinction,[14] proportionality, precautions in attacks[15] and a number of other IHL principles and rules. Thus responding to the increase of war crimes committed in armed conflicts, the development of ICL has contributed and is contributing to filling gaps existing in IHL instruments, whilst at the same time supporting the process of osmosis from IAC to NIAC law which is one of the most important defining elements of contemporary IHL.

International tribunals have also played a crucial role in asserting the continuing application of IHRL in armed conflict. Since the two famous advisory opinions of the International Court of Justice (ICJ) on the legality of nuclear weapons (1996) and on the construction of the wall by Israel in the Palestinian territory (2004), it is recognized that the protection of human rights norms does not cease in time of armed conflict and the relationship between IHRL and IHL is defined as one of *lex generalis* / *lex specialis*.[16] Thus the European Court of Human Rights (ECtHR) delivered judgments finding violations of the European Convention of Human Rights (ECHR) in connection with the conflict in Chechnya, including killing and injuries to civilians, destruction of homes and property, use of landmines, torture and inhuman conditions of detention.[17] The Inter-American Human Rights control and judicial mechanisms also offer important examples of application and interpretation of IHL norms.[18]

---

[13] According to the famous ICTY decision in the Tadić case "an armed conflict exists whenever there is a resort to armed force between States or protracted armed violence between governmental authorities and organized armed groups or between such groups within a State" (IT-94-1, 2 October 1995, para. 70).

[14] IT-95-16-T, Kupreškić, 14 January 2000, para. 521; IT-01-47-AR73.3, Hadžihasanović, 11 March 2005, para. 30.

[15] IT-98-29-T, Galić, 5 December 2003, para. 58

[16] ICJ, Legality of the Threat or Use of Nuclear Weapons, Advisory Opinion of 8th July 1996, para. 25; ICJ, Legal Consequences of the Construction of a Wall in the Occupied Palestinian Territory, Advisory Opinion of 9th July 2004, para. 106. See Sivakumaran, *The Law of Non-International Armed Conflict*, supra n. 5 at pp. 88-93.

[17] The ECtHR applied Articles 2 (right to life), 3 (prohibition of torture and inhuman or degrading treatment), 5 (right to liberty and security), 8 (right to respect for private and family life), 13 (right to an effective remedy) and 14 (prohibition of discrimination) of the ECHR and Article 1 (protection of property) of Protocol No. 1 to the ECHR. See www.echr.coe.int/Documents/FS_Armed_conflicts_ENG.pdf, 19th September 2019 at 10-12.

[18] See E.J. Buis, 'The Implementation of International Humanitarian Law by Human Rights Courts: the Example of the Inter-American Human Rights System' in R. Arnold and N. Quénivet, *International Humanitarian Law and Human Rights Law*, Brill-Nijhoff, Leiden, 2008, pp. 269-293 at pp. 277-292.

The role of legal thought

Elaborating legal analyses of the changing nature of armed conflicts was not the prerogative of the sole judicial practice. Legal scholarship has produced substantial research related to the contemporary forms of armed conflicts and the applicable substantive law. Leading academics, military experts and institutions have expressed theoretical orientations relating to IHL both individually and through collective works serving as references for the specialist, the diplomat and the judge. Important manuals and studies have covered the law of naval warfare as well as the law of air and missile warfare,[19] the law applicable to international operations,[20] the legal regulation of cyber warfare[21] and a number of other areas. The contributions of the ICRC deserve a special mention, particularly its fundamental Study on Customary International Humanitarian Law stating that most of the customary IHL rules are applicable in international as well as non-international armed conflicts and are binding on both sides to a conflict.[22] The analysis of international practice has enriched our knowledge of IHL and the ability of those who operate in the field to implement its rules in IACs as well as in NIACs.

International legal scholars have delved into the concept of armed conflict paying special attention to the different types of NIAC and to IHL applicable thereto. There has been academic debate about whether IHL is still based on a binomial IAC / NIAC system or – as some scholars argue – the progressive convergence of IAC *jus in bello* and the law of NIAC is leading to a blending of the law regulating the two types of armed conflict.[23] In the case of so-called spill over conflicts it is generally

---

[19] See the *San Remo Manual on International Law Applicable to Armed Conflicts at Sea*, 1994 and the *HPCR Manual on International Law Applicable to Air and Missile Warfare*, 2013.

[20] See T. Gill and D. Fleck D, Ed.s *The Handbook of the International Law of Military Operations*, 2nd ed., Oxford University Press, Oxford, 2015 and the *Leuven Manual on the International Law Applicable to Peace Operations*, Cambridge University Press, Cambridge, 2017.

[21] See the *Tallinn Manual on International Law Applicable to Cyber Warfare*, 2013.

[22] See J-M. Henckaerts and L. Doswald-Beck, *Customary International Humanitarian Law*, 2 Volumes, Cambridge University Press, 2005. The Study is updated through the Customary IHL Database available at https://ihl-databases.icrc.org/customary-ihl/eng/docs/home.

[23] R. Bartels 'Timelines, borderlines and conflicts. The historical evolution of the legal divide between international and non-international armed conflicts,' IRRC Vol. 91 (2009) No. 873, reviews the different opinions of scholars (pp. 40-41) concluding that at present the distinction between the two types of conflict still forms part of positive law (p. 67).

recognized that the expansion of an internal fighting into the territory of a foreign state would not change the nature of the conflict and would remain fully subject to the law of NIAC.[24] More problems arise with military intervention due to the variety of cases that may occur. Most recent history has shown states fighting NSAGs which operate from the territory of a foreign state, intervening militarily in the territory of that state, with or without its consent; states or multinational coalitions intervening in support of the official government of another state against local insurgents, but also to support insurgents against the incumbent government, and even NSAGs supporting the government of a state fighting against other NSAGs. As a consequence, several and distinct armed conflicts may occur in the same territory, either simultaneously or consecutively; in some situations, an internal conflict may become international but an inter-state conflict may also transition to an internal conflict due to changing circumstances.[25] In the majority view, the existing binary IHL framework is still adequate for the purpose of regulating contemporary armed conflicts, with the consequence that each situation should be classified according to the thresholds of armed conflict and the applicable law should be determined on a case by case basis – with all the difficulties this entails for those operating on the field.[26]

Concluding remarks

In the decades following the adoption of the Geneva Conventions and Protocols the emergence of new types of armed conflicts has given rise to sensitive questions about which IHL rules apply in each different situation and whether the traditional binary paradigm IAC / NIAC is still an

[24] See Dinstein, *Non-International Armed Conflicts* supra n. 8, p. 25.

[25] See S. Vité, 'Typology of armed conflicts in international humanitarian law: legal concepts and actual situations,' IRRC Vol. 91 (2009) No. 873 pp. 69-94 at pp. 83-93; M. Milanovic and V. Hadzi-Vidanovic, 'A taxonomy of Armed Conflict' in N. White and C. Henderson, *Research Handbook on International Conflict and Security Law*, Edward Elgar Publishing, Cheltenham, 2013 pp. 256-314 at pp. 291-298; T. Ferraro, 'The ICRC's legal position on the notion of armed conflict involving foreign intervention and on determining the IHL applicable to this type of conflict,' IRRC Vol. 97 (2015) No. 900 pp. 1227-1252 at pp. 1240-1250.

[26] See Bartels, 'Timelines, borderlines and conflicts' supra n. 24. See also Vité, 'Typology of armed conflicts' supra n. 26 at p. 86 and Ferraro 'The ICRC's legal position' supra n. 26 at p. 1229.

adequate model to effectively reflect the changing nature of contemporary armed conflict.

From the normative point of view, the two parts of IHL remain clearly separate: on the one hand, the GCs and AP I governing IAC; on the other hand, CA 3 and AP II being applicable to NIAC. However, as said above (4.3), several other treaties have expressly extended their scope to NIAC thus establishing a legal bridge between the two sets of rules, particularly those related to the protection of cultural property and the use of weapons. The interpretive-analytic activity as developed by the case law of international criminal tribunals also argues in favour of a progressive harmonization between the two parts of IHL.

Clearly, sensible divergences between the law of IAC and the law of NIAC still persist, the most relevant being the treatment of persons deprived of their liberty, the regulation of occupied territory, and the obligations of states not involved in the conflict. While in IACs combatants are entitled to POW status and the laws of neutrality and belligerent occupation are applicable, these areas are not covered by the law of NIACs. Another critical issue deserving discussion on both formal and substantive grounds relates to equality between the parties of a NIAC. While the legal basis of the obligation of NSAGs to respect IHL can be found in treaties and in customary law, their capability and willingness to abide by the existing rules are another matter altogether. For this reason, engaging NSAGs in implementing IHL is one of the most demanding challenges facing IHL today.

Although the legal qualification of an armed confrontation is complicate, especially for those operating on the field in a territory where several and distinct conflicts may occur, this certainly must not prevent the quest for the best protection of victims by all actors involved. It should be taken into account that HRL, general public international law and even soft law (e.g. for situations of detention) can provide suitable rules to be applied in the diverse landscape of contemporary armed conflicts.

A legacy of responding to new means and methods of warfare: the regulation of new weapons under international law

Hitoshi NASU
Professor of International Law, Exeter University

It is often said that the legal regulation of weapons is lagging behind technological developments. In this presentation, I challenge this proposition by advancing three reasons why we should exercise caution against exaggerating regulatory problems in relation to new weapons.

In a nutshell, there are three reasons: (1) the legacy of weapons law principles; (2) the importance of adequately understanding the characteristics, potential and limitation of new technologies; and (3) the power of non-legal forms of regulation.

First of all, weapons law 101. There are two general principles under customary international law prohibiting, first, the employment of arms, projectiles or material 'calculated to or of a nature to cause superfluous injury or unnecessary suffering', and second, the use of weapons that indiscriminately affect both lawful targets and civilians. These two principles address two different humanitarian concerns. The first principle – superfluous injury or unnecessary suffering – aims to limit the degree of injury or suffering inflicted upon lawful targets, relative to the military necessity underlying the choice of a particular weapon. The second principle – indiscriminate weapons – is, on the other hand, designed to protect civilians from the effects of the weapon.

Various weapons treaties address these two humanitarian concerns with reference to specific types of weapon. For example, the ban on certain types of explosive projectiles, expanding bullets, and non-detectable fragments are specific manifestations of superfluous injury or unnecessary suffering as agreed by states. Whereas the ban on anti-personnel mines and cluster munitions was driven more by concerns about their indiscriminate effect on civilians. The legal regulation of these weapons is not necessarily considered as failure to keep up with technological developments. Indeed, the 1868 St Petersburg Declaration and the 1899 Hague Declaration III were both adopted to prohibit certain types of explosive projectiles and expanding bullets respectively, as these new weapons emerged for use in

the battlefield. Non-detectable fragments and blinding lasers were both banned before these technologies became operationalised in combat.

The absence of specific treaty prohibition does not mean that new weapons are unregulated under international law. Their lawfulness must still be assessed in light of the general principles explained earlier and the new weapon may well be considered to cause superfluous injury or unnecessary suffering or be indiscriminate in nature. Difficulties, however, arise from the application of these principles. The intended injury or suffering is considered superfluous or unnecessary only against the underlying military value attached to the new weapon. Whether the new weapon is indiscriminate in nature or not may well be situation dependent. These general principles are, in essence, flexible but elusive in application.

The legality of any new weapon can be a subject of debate, but even in such situations the general principles can provide normative guidance for the debate. As explained earlier, the two general principles articulate two different humanitarian concerns – one regarding the degree of injury or suffering inflicted upon lawful targets, and the other regarding indiscriminate effect upon civilians. The regulatory debate about any new weapon should, at least, proceed with clarification of which one of these humanitarian concerns is raised and is to be addressed. For example, there is no point reassuring that lethal autonomous weapons are capable of discriminating lawful targets from civilians for people who are concerned about the idea of machines killing a person. Likewise, emphasising or requiring human control in the use of lethal autonomous weapons does not necessarily help address the indiscriminate effect of the weapon system.

The second reason why we should not overstate regulatory problems with new weapons is that our understanding of emerging technologies tends to be limited. This is because the development of a new technology is not a sequential process, but rather involves a complex web of scientific findings and technological breakthroughs. Once developed, the technology is further refined for improvement and sophistication often with various tailored applications. Consider, for example, how technology evolved for vehicles, aircraft and smartphones, just to name a few.

A poor understanding of the characteristics of any new technology and its potential applications and limitations often results in fearmongering campaigns, exaggerating the risks and dangers technology might pose. The ‘fear’ factor is an inevitable human condition as the instinctive and primitive response to unknowns. However, any regulatory attempt driven by fear is destined to be short-lived, as has been proved by the failure to restrict aerial warfare with the use of balloons for discharging projectiles

and explosives. The Declaration was adopted in 1899 to set a moratorium on the use of balloons for launching attacks, but soon later, major military powers realised the strategic advantage that aerial warfare would bring to the battlefield and refused to agree with the renewal of the restriction.

It is always advised that experts, including lawyers, exercise their due diligence by developing adequate understanding of the subject matter before drawing any conclusions, and do not make any assumption about technological capabilities and functional parameters. This means that it necessarily takes time to develop sufficient understanding of the characteristics of a new technology as the basis for adequate assessment regarding the need for new regulation and the ways in which its applications should or can be regulated.

The third reason is the power of non-legal forms of regulation. It is not legal considerations that direct weapons development programmes. Rather, constraints on weapons development are derived from other factors, such as strategic and political considerations, technological feasibility, financial costs, and existing military infrastructures. Legal regulation is only one form of controlling the means and method of warfare, and quite often, not a decisive one. Legal regulation is not necessarily lagging behind when other forms of regulation are available to regulate the development and use of new weapons.

Consider, for example, the constraint on the use of depleted uranium (DU) weapons. While DU munitions are not prohibited, there was strong public reaction against the use of DU munitions in the aftermath of the two Gulf Wars and military operations in Bosnia, Kosovo and Afghanistan. Despite disputed scientific evidence regarding its health and environmental effects, there has since been a significant reduction in the stockpiling and use of DU munitions in many countries. Nuclear weapons is another example. There has been no single instance of nuclear launches in combat since Hiroshima and Nagasaki, despite the fact that nuclear weapon States have been persistent in denying the illegality of nuclear weapons.

For effective regulation of new weapons, the mere adoption of a specific treaty is not as important as the process leading up to it and building public pressure to raise political costs associated with the development and use of the weapon. In this respect, we should acknowledge the significant power of public campaign as a non-legal form of regulation constraining the development and use of a new weapon. Because of this power, public campaigns should be employed wisely, not blindly or driven by fear, in light of realistic assessment of technological capabilities and their potential role in military affairs.

The oft-quoted observation that the legal regulation of weapons is lagging behind technological developments appears to be based on a myopic view, focusing solely on the specific treaty prohibiting or restricting the use of a particular weapon. It is my submission that we should adopt a broader perspective to the regulation of new weapons, with careful and evidence-based assessment of technological characteristics, while seeking guidance from the general principles of international humanitarian law.

From land, to sea, to air – from the trenches to the city: international humanitarian law and the changing realities in the conduct of hostilities during the past century

Gloria GAGGIOLI
Former Judge Advocate General, Canadian Armed Forces

The waging of wars is in constant evolution. My co-panelists have addressed this evolution in relation to the nature of armed conflicts and the nature of weapons. I will focus on: first, the multiplication of warfare domains (land, sea, air, but also cyber or space) and, second, I will address the urbanization of warfare, i.e. the fact that combats are not taking place in trenches or on a battlefield in open country but right in the middle of cities or densely populated areas. I will not elaborate on the historical evolutions (except for recalling briefly the when and why of these evolutions), but rather focus on the humanitarian and legal challenges pertaining to these two types of warfare evolution.

From land, to sea, to air…

a) The multiplication of warfare domains in past centuries

Contrary to what the title of my presentation seems to induce, the emergence of new warfare domains does not make the old ones disappear. They rather pile up. In the military context, the phrase "multidimensional" or "multi-domain" is increasingly used to describe the expansion and interrelated character of these various domains.[1]

The emergence of new warfare domains is not a novelty. While land warfare is as old as mankind, naval warfare[2] can be traced back to more

[1] See e.g. *Multi-Domain Battle: Evolution of Combined Arms for the 21st Century 2025/2040,* Dec. 2017. Available at: www.tradoc.army.mil/Portals/14/Documents/MDB_Evolutionfor21st%20(1).pdf.

[2] "Naval warfare" is the term used to denote "the tactics of military operations conducted on, under, or over the sea". See Encyclopedia Britannica, online: www.britannica.com/search?query=naval+warfare.

than 3,000 years ago (e.g. the Battle of the Nile Delta, 1175BC).[3] Naval warfare has witnessed considerable evolutions in the mid-19th century. By 1914 technological innovations produced far more powerful and capable warships than those of the Age of Sail (17th century - mid-19th century) and submarines were widely used during World War I (1914–1918). Instead, the military use of airpower (other than kites and balloons) is more recent. It was only with the First World War that airplanes (heavier-than-air aircraft) were used in support of the army on the ground and the navy on the surface.[4] The Second World War, and the "total war" ideology that prevailed at the time, fostered the "strategic" use of airpower with the political objective to undermine the morale of the enemy through intensive bombings.[5] The use of airpower continues to experience evolutions with, for instance, the invention and military use of drones (first drone strike in 2001).[6] In our century, cyberspace is often described as a new warfare domain; although some consider that "cyber" is more about means and methods of warfare because the effects are felt on land, sea, air.[7] In operational terms, it is nevertheless still useful to describe cyberspace as a warfare domain. Space warfare, although it never materialized, is currently been discussed by experts in order to articulate the rules that would be applicable to military space operations.[8]

As for weapons, the emergence of new warfare domains goes hand in hand with the discovery of new technologies. Historians, such as Eric Germain, tend to consider that the First World War played a pivotal role in the emergence of the multidimensional battlespace (or the "the

[3] Rolf Fabricius Warming, "An Introduction to Hand-to-Hand Combat at Sea – General Characteristics and Shipborne Technologies from c. 1201 BCE to 1600 CE", in Johan Rönnby (ed.), *On War On Board – Archeological and Historical Perspectives on early modern maritime violence and warfare*, Södertörns högskola, Stockholm, 2019, p. 108.

[4] Interview with Richard Overy, *International Review of the Red Cross*, vol. 97, n° 900, 2015, at 969.

[5] Ibid, 971.

[6] Arthur Holland Michel, "How rogue techies armed the predator, almost stopped 9/11, and accidentally invented remote war", *Wired,* 17th Dec 2015. Available at: www.wired.com/2015/12/how-rogue-techies-armed-the-predator-almost-stopped-911-and-accidentally-invented-remote-war/.

[7] ICRC, *International Humanitarian Law and the Challenges of Contemporary Armed Conflicts*, Report submitted to the 31st International Conference of the Red Cross and Red Crescent, 2011, at 36. Available at: www.icrc.org/en/doc/assets/files/red-cross-crescent-movement/31st-international-conference/31-int-conference-ihl-challenges-report-11-5-1-2-en.pdf.

[8] See further below.

globalization of the battlefield" to quote him) as we know it today.[9] Interestingly, he also suggests that the preservation instinct and traumatic events such as the carnages in the trenches have also motivated the human being to extend the range of weapons and to discover new warfare domains in order to get away from the enemy while still being able to inflict harm.[10] This is certainly true for aerial warfare and may also be an explanatory factor for cyber warfare, for instance, where the attacker becomes not only invisible but potentially anonymous.

b) What is the humanitarian impact of the multiplication of warfare domains?

On that basis, a first question we might ask ourselves is whether the emergence of new warfare domains has humanitarian impacts and, in particular, whether it leads to an increase in deadliness and/ or in the ratio of civilian-military casualties. We might assume an increase in deadliness. Intuitively, it seems logical to expect a potential increase in deadliness (among uniformed and civilian personnel alike) because of the mere fact that there is a multiplication of battlespaces (land, air, sea etc.). For instance, it is well-known that the two world wars, which coincided with a qualitative leap in warfare technologies, were particularly deadly. However, empirical data tends to show that today's wars are not necessarily more deadly than in the past. "Our World in Data", a collaborative effort between researchers at Oxford University, has gathered and analysed data on past conflicts that tend to demonstrate that the absolute number of war deaths is declining since 1945 and that, more generally contemporary conflicts are not necessarily deadlier than in the past (see figure 1).[11]

Instead, there is almost unanimity among experts that the ratio civilian-military casualties has steadily increased. For instance, after having conducted a careful study of existing statistics on the matter, Prof. Valerie Epps concludes that "it seems more than fair to conclude that since the turn of the twentieth century, civilian deaths have outnumbered military deaths in nearly all wars."[12] Now, the correlation between the multiplication of

[9] Eric Germain, "Out of sight, out of reach: Moral Issues in the Globalization of the Battlefield", *IRRC,* vol. 97, n°900, 2015, p. 1066.

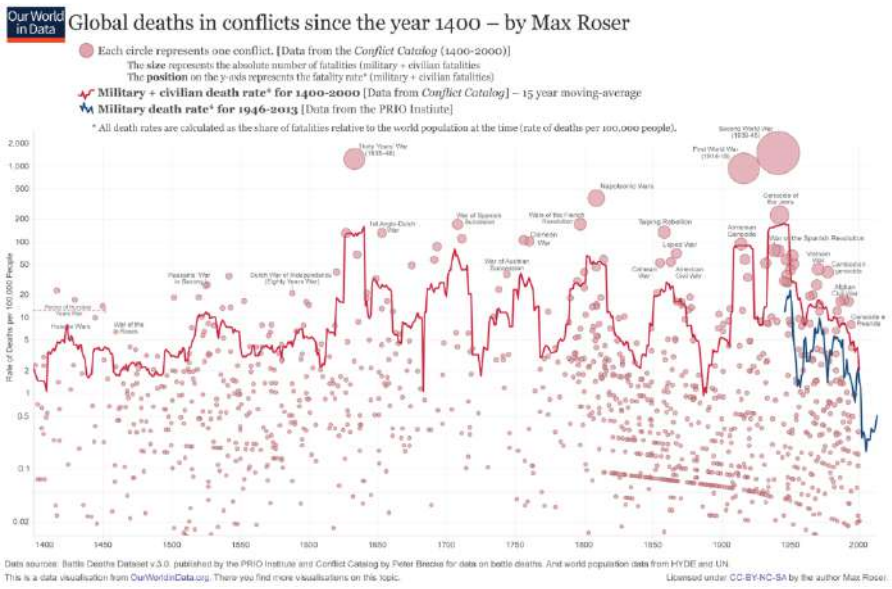
[10] Ibid, p. 1068.

[11] Max Roser, *War and Peace*, 2020. Published online at OurWorldInData.org. Retrieved from: https://ourworldindata.org/war-and-peace [Online Resource]

[12] Valerie Epps, "Civilian Casualties in Modern Warfare, The Death of the Collateral Damage Rule", *Georgia Journal of International and Comparative Law*, 2013, at 329.

warfare domain and increase in the civilian-military casualties' ratio is difficult to establish. It seems nevertheless fair to argue, for instance, that the emergence and progresses in airfare allowed the strategic use of "airpower", i.e. the attempted "killing of cities" during WWII[13] and, therefore, increases in civilian casualties. In the same vein, the development of naval warfare and invention of submarines allowed torpedo attacks on neutral vessels, merchant ships and hospital ships during WWII, which in turn led to an increase in civilian casualties. While, in the past, new warfare domain went hand in hand with an increase of casualties among civilians, this is not an unavoidable consequence. Depending on how wars are being fought in new warfare domains, deaths among civilians may be avoided; in particular, when new technologies allow for more precise means and methods of warfare. In brief, a nuanced approach is needed. New warfare domains have obviously the potential to increase deadliness and/or civilian-military casualties' ratio, but this is not inevitable.

Figure 1



[13] Interview with Richard Overy, *International Review of the Red Cross*, vol. 97, n° 900, 2015, at 972 and ff.

c) What are the legal challenges pertaining to the discovery of new combat fields?

Another question is whether new warfare domains give rise to specific and/or new legal challenges. Interestingly, a cursory analysis of the discussions in the international community in relation to new warfare domains throughout history tends to demonstrate that rather similar legal issues emerged irrespective of the warfare domain in question. Be it for cyber or aerial warfare, when these domains emerged, States and experts asked themselves the same questions in broad terms, such as: is the law as it exists sufficient/adequate to regulate the new warfare domain? Should new specific rules be adopted?

In this respect, we can notice a fairly systematic lack of appetite by States to regulate new warfare domains, even after wars have demonstrated that current rules are obsolete or insufficient. In the context of naval warfare**,** for instance, although eight Hague Conventions were adopted in 1907 to regulate this domain, they became rather soon outdated with technological advances during the WWI and they were never replaced with a more recent treaty. There were attempts to develop new rules (e.g. 1909 London Declaration), but these merely resulted in the adoption of a single provision in 1930 stating that "submarines must conform to the rules of international law to which surface vessels are subject".[14] And while the Second Geneva Convention of 1949 deals with wounded, sick and shipwrecked at sea, it does not address the conduct of hostilities. As for Additional Protocol I to the four Geneva Conventions, it applies to sea warfare, but only when it may affect civilians on land.[15] Regarding air warfare, the lack of treaty provisions governing specifically air warfare is well-known. The Hague Declaration (XIV) of 1907[16] prohibited the discharge of projectiles and explosives from balloons or other similar new methods; but at the time air technology was not sufficiently advanced to permit the precise targeting of objectives to be destroyed. Nothing specific to air warfare is to be found in the 1949 Geneva Conventions and Additional Protocol I applies to air warfare, but again only insofar as it produces effects on land.[17] Lastly, no specific treaty rules have so far been adopted in relation to cyber or outer space warfare.

---

[14] Art. 22 of the London Treaty on Limitation and Reduction of Naval Armaments (1930) and the Procès-Verbal on Submarine Warfare of the Treaty of London.

[15] See Art. 49, para. 3, of Additional Protocol I.

[16] Declaration (XIV) Prohibiting the Discharge of Projectiles and Explosives from Balloons, The Hague, 18th October 1907.

[17] See Art. 49, para. 3, of Additional Protocol I.

The response of the international community to these gaps in norm-settings has often been to draft non-binding expert documents in order to: 1) re-state that general international humanitarian law (IHL) rules and principles, in particular in relation to the conduct of hostilities (distinction, proportionality and precautions) exist and apply to the new warfare domain; and 2) articulate specifically those rules in the new warfare domain. The 1994 *San Remo Manual on International Law applicable to Naval Warfare*,[18] the 2009 *Manual on International Law applicable to Air and Missile Warfare*,[19] the 2013 *Tallinn Manual on the International Law applicable to Cyber Warfar*e and 2017 *Tallinn 2.0 on the International Law applicable to Cyber Operations*,[20] the *Woomera Manual for Military Space Operations* (expected date of publication 2021)[21] all have in common this broad objective to fill gaps, or at least to clarify what the law allegedly is (or should be) in these areas.

These manuals comport many advantages.[22] They may be used by courts, tribunals, organizations; they may influence State practice and thus

[18] Available at: https://ihl-databases.icrc.org/ihl/INTRO/560. The San Remo Manual was developed by a group of legal and naval experts in the context of a series of Round Tables convened by the International Institute of Humanitarian Law. It is viewed by some as the modern equivalent to the *Oxford Manual on the Laws of Naval War Governing the Relations Between Belligerents* adopted by the Institute of International Law in 1913. It is nowadays largely accepted as reflecting customary IHL norms.

[19] Available at: https://reliefweb.int/sites/reliefweb.int/files/resources/8B2E79FC145BFB3D492576E00021ED34-HPCR-may2009.pdf. This Manual was developed by an international group of experts convened on several occasions by the Program on Humanitarian Policy and Conflict Research at Harvard University (HPCR). See also the 1923 Hague Rules of Air Warfare, a draft treaty prepared by a Commission of jurists that has never been adopted by States. Available at: https://ihl-databases.icrc.org/ihl/INTRO/275. This document focused on the prohibition of targeting undefended localities on land but did not prohibit the indiscriminate attacks of defended cities. It was considered by some at the time as reflecting customary law.

[20] See: https://ccdcoe.org/research/tallinn-manual/. The Tallinn Manual was developed by an international group of legal scholars and practitioners and convened by the NATO Cooperative Cyber Defence Centre of Excellence.

[21] See: https://law.adelaide.edu.au/woomera/system/files/docs/Woomera%20Manual.pdf. The Woomera Manual is being developed by a group of legal experts specialized in the fields of international space law, international law on the use of force and IHL, together with technical experts. The project is funded by the University of Adelaide, the University of Exeter, the University of Nebraska College of Law and the University of New South Wales in Canberra.

[22] For an appraisal of the impact and a critique of international operational manuals, see, e.g.: Dale Stephens and Melissa De Zwart, "The Manual of International Law applicable to Military Uses of Outer Space (MILAMOS), *RUMLAE Research Paper* 2017, pp. 3-5. See also the keynote speech given by Professor Yoram Dinstein in the context of the 2019 Sanremo Round Table.

potentially contribute to the emergence of customary norms; their drafting may be relatively rapid, flexible and less politicized; and they leave room for progressive interpretations. But their downsides are also real. They are not legally binding *per se* and, therefore, cannot be easily used to hold States accountable. They lead to legal uncertainties as to what is actually legally required. They are not State-driven and may, therefore, lack ownership among those who have to apply the rules. Experts who develop these instruments often come from Western countries and lack representativeness. The methodology used to develop these instruments is sometimes unclear and not transparent. As a result, these manuals are certainly useful, or even needed given the lack of appetite by States to develop new instruments, but they remain second-best solutions. As aptly noted by Dale Stephens and Melissa De Zwart: "For the better or worse, this is the 'age of the manual' and their continued drafting by private experts signifies a need to fill gaps that have been unwittingly left by States as new technologies and capabilities emerge. The status of these International Operational Law Manuals is self-declared to be non-binding and yet they do seem to nonetheless attract significant normative traction."[23]

The legal vacuum or at least the lack of clarity in the rules to be applied in the context of new warfare domains has, potentially in turn, a humanitarian impact. If States were able to agree on clear and specific rules when new warfare domains emerge, this would contribute to prevent an increase in casualties among the civilian population. The existence and clarity of rules have indeed a preventive effect. At least, this is what lawyers tend to believe and hope for.

From the trenches to the cities…

a) The urbanization of warfare: when and why?

In parallel to the emergence of new warfare domains, the traditional distinction between the "front" and the "rear" has disappeared.[24] In the mid-20th century, combats have moved away from the Great War's trenches and became increasingly fought within the cities. Precursors of urban warfare as we know it today can be traced back to the 1930s with, for instance, the Spanish Civil War (36-39).[25] I have already mentioned the

[23] Dale Stephens and Melissa De Zwart, ibid, p. 6.

[24] Vincent Bernard, "Editorial: War in Cities, the Spectre of Total War", *IRRC: War in Cities*, vol. 98, n° 901, April 2016, p. 6.

[25] Ibid, p. 2.

attempted "killing of cities" during the Second World War, which was tightly linked with the concept of total war.[26]

This urbanization of warfare became even more marked with the Cold War and the related multiplication of non-international armed conflicts and asymmetric warfare. In the words of the British General Sir Rupert Smith, wars became "wars amongst the people".[27] Organized non-state actors blend in with the cities' civilian population and bring combat right at the heart of densely populated areas. As highlighted by Vincent Bernard, "[a]rmed groups are sometimes born in cities, or they may hide in cities to benefit from the terrain: drawing the enemy into terrain that gives you an offensive or defensive advantage is a basic tactical ploy, and fighting in a city allows armed groups to make up for their relative weakness in these 'asymmetric warfare' situations."[28] This trend is even more pronounced in the context of the so-called war on terror, where organized armed groups labelled as terrorists not only conduct hostilities within cities but also direct their military operations against civilians and civilian objects. "Because cities are highly symbolic, they are also the preferred target of terror attacks, recent examples being New York, Mumbai, Paris and Nairobi."[29] Even siege warfare (a medieval method of warfare that seemed to be forgotten) is back, as evidenced by recent practices in Syria or Iraq.[30]

The inevitable trend of urbanization and multiplication of "megacities" give rise to legitimate concerns to protect civilians against the effects of hostilities in urban settings. According to the United Nations (UN) 2018 Revision of the World Urbanization Prospects report, 55% of the world's population resides in urban areas in 2018; and one in eight of the world's urban dwellers live in 33 megacities with more than 10 million inhabitants.[31] Additionally, the UN projects that 68% of the world's population will reside in urban settings in 2050 (while the figure was only 30% in 1950).

---

[26] See above, p.3.

[27] Interview with General Sir Rupert Smith, *IRRC: Methods of Warfare,* Vol.88, n° 864, Dec. 2006, p. 719.

[28] Bernard, above n. 24, p. 4.

[29] Ibid., p. 3.

[30] See e.g. Gloria Gaggioli, "Besieging cities and humanitarian access: how to accommodate humanitarian needs, legal obligations and operational constraints?", *Proceedings of the 20th Bruges Colloquium,* forthcoming. See also: G. Gaggioli, "Are Sieges Prohibited Under Contemporary IHL?", *EJIL: Talk!,* 2019. Available at: www.ejiltalk.org/joint-blog-series-on-international-law-and-armed-conflict-are-sieges-prohibited-under-contemporary-ihl/.

[31] United Nations, Department of Economic and Social Affairs, Population Division, *World Urbanization Prospects: The 2018 Revision,* New York, United Nations, 2019 *(ST/ESA/SER.A/420)*, p. xix. Available at: https://population.un.org/wup/Publications/Files/WUP2018-Report.pdf.

b) The humanitarian impacts of the urbanization of warfare

A number of experts have highlighted that the urbanization of warfare has devastating humanitarian impacts. Fighting in densely populated areas automatically implies that civilians are increasingly affected by warfare. Aleppo, Gaza, Mogadishu, Mosul and Sana'a constitute just a few contemporary examples of the urbanization of warfare.[32] In 2016, ICRC President Peter Maurer was of the view that Aleppo was experiencing "one of the most devastating urban conflicts in modern times".[33]

Civilians may be targeted or killed incidentally as a result of fighting, mines or improvised explosive devices. Cultural heritage is being destroyed. Civilians are being displaced and relocated in overcrowded refugee or IDP camps. Urban services disintegrate and deprive the populations from power, water and food supplies. Health care becomes poor or non-existent. These situations give rise to immense challenges for humanitarian organizations, in terms of access, security for its own staff and/or to ensure the evacuation of the population. The urbanization of warfare also has long-term impacts on the economy, education and healthcare systems.

c) The legal challenges pertaining to the urbanization of warfare

In the context of the urbanization of warfare, the main legal issue does not concern the question whether new laws should be adopted. If basic IHL rules – such as prohibition of indiscriminate attacks – were respected in urban warfare, this would already spare numerous civilians. A first issue is to ensure respect for existing rules. It relates to implementation. Another issue is how to interpret current rules and principles in the context of the urbanization of warfare. Is there room for clarification or even evolutionary interpretation within contemporary IHL and, if yes, how much?

For instance, how to assess the principle of proportionality in urban warfare.[34] Should so-called "reverberating effects" or "knock-on effects" be

---

[32] See e.g. Margarita Konaev and John Spencer, The Era of Urban Warfare is Already Here, E-Note, 21st March 2008. Available at: www.fpri.org/article/2018/03/the-era-of-urban-warfare-is-already-here/.

[33] ICRC News Release. Available at: www.icrc.org/en/document/syria-news-cities-aleppo-one-most-devastating-urban-conflicts.

[34] See in this sense, *International Humanitarian Law and the Challenges of Contemporary Armed Conflicts, Recommitting to Protection in Armed Conflict on the 70th Anniversary of the Geneva Conventions - Report*, Geneva, ICRC, 2019, p. 7. See also: ICRC, *The Principle of Proportionality in the Rules Governing the Conduct of Hostilities under International Humanitarian Law*, 2018; available at www.icrc.org/en/download/file/79184/4358_002_expert_meeting_report_web_1.pdf.

taken into account to assess the "expected incidental civilian casualties and damage to civilian objects as required under the IHL principles of proportionality and precautions"? For instance, the destruction of a transformer may be expected to shut down a whole hospital. Should the destruction of the hospital be considered as part of the expected incidental civilian casualties?

In the ICRC's view, reverberating effects that are foreseeable for a reasonably well-informed military commander need to be included in the assessment of expected incidental civilian casualties.[35] Taking into account reverberating effects in the proportionality assessment is particularly important in the context of the use of explosive weapons in populated areas, which often disrupts the functioning of essential services.[36] In a report on the principle of proportionality produced by Chatham House, a similar analysis is provided.[37] Others consider that reverberating effects should be left out because it is impossible in practice to consider these effects as "expected" given the number of variables outside the attacker's control that may influence the outcome.[38] Still others have attempted to devise criteria to limit the type of indirect effects that should be taken into account (e.g. only those which are "likely" or "almost inevitable" or which have a "close nexus" with attack).[39]

Taking into account reverberating or knock-on effects is certainly an evolutive interpretation of the principle of proportionality under IHL. The mere notion of reverberating effects or knock-on effects was (and still is) absent from most IHL scholarly textbooks and military manuals. But this novel interpretation is certainly necessary from a humanitarian perspective, correct from a legal perspective – since nothing in the principle of proportionality limits expected "collateral damages" to those which are directly caused by the attack[40] – and realistic today given the evolutions in

---

[35] ICRC Q&A on the issue of explosive weapons in populated areas, *IRRC,* vol. 98, n° 1, 2016, p. 104. See also: *International Humanitarian Law and the Challenges of Contemporary Armed Conflicts*, 2015, 32IC/15/11, pp. 52-53; available at www.icrc.org/en/download/file/15061/32ic-report-on-ihl-and-challenges-of-armed-conflicts.pdf; ICRC Challenges Report 2019, p. 9.

[36] ICRC, *The Principle of Proportionality* …, above n. 34, p. 45.

[37] Emanuela-Chiara Gillard, *Proportionality in the Conduct of Hostilities: the Incidental Harm Side of the Assessment,* Chatham House, December 2018, pp. 18-20.

[38] ICRC, *The Principle of Proportionality* …, above n. 34, p. 45 and ff. (presenting different views on the matter)

[39] Ibid.

[40] In the proportionality assessment, one has to compare "expected" incidental civilian damages with the "concrete and direct military advantage anticipated". While the anticipated military advantage must be direct, the incidental civilian damages must merely be expected (no direct causation is required). See Article 51§5b) API.

technologies and amount of intelligence that can be gathered by belligerents (e.g. through drones). It is, therefore, to be hoped that belligerents, and military commanders in particular, will seek advice from not only legal advisers but also engineers and architects in the years to come.

Conclusion

The evolution of warfare has taken many different forms. It is sometimes the result of new technologies (e.g. invention of submarines or airplanes). At other times, it is simply the result of socio-political evolutions (e.g. urbanization linked to industrialization, political/economic crisis, natural population increase, social changes/lifestyle changes).[41] Each of these evolutions give rise to particular humanitarian challenges and have the potential for increased casualties particularly amongst the civilian population and long-term disruptions of societies.

The evolution of warfare equally gives rise to legal challenges. In the context of new warfare domains, the issue whether the law should be developed is recurrent. States are generally reluctant to develop new rules, maybe because they do not wish to be further constrained by the law, or maybe because they consider existing rules as sufficient. In any case, there is a wide consensus that principles of IHL, such as the principles of distinction, proportionality and precautions equally apply on land, air, at sea, in the cyberspace or potentially in the outer space. In the absence of specific treaty rules for naval warfare, air warfare, cyber and outer space, experts develop soft law instruments in the shape of international operational manuals, but these are only second-best solutions.

In the context of the urbanization of warfare, the main legal issues pertain to implementation and the interpretation to be given to key international law rules and principles of IHL, such as the principle of proportionality. In my view, evolutions in the realities of armed conflicts must go hand in hand with evolutionary interpretations of IHL and an actualization of the reading of IHL provisions. IHL should not be seen as a frozen body of law that is incapable of dealing with the evolution of warfare. On the contrary, IHL rules must be interpreted in an evolutionary manner in order to cope with these evolutions, while always having in mind the indispensable balance between the principles of military necessity and humanity.

[41] See e.g. www.theclassroom.com/what-are-the-causes-of-urbanization-in-poor-countries-13660201.html.

II. IHL and the challenges related to cyber warfare

Casualties caused through computer network attacks: the potential human costs of cyber warfare

Marina KROTOFIL
Senior Security Engineer, BASF

Summary

Industrial Control Systems (ICS) threat landscape has changed dramatically over the past few years. New threats have emerged to challenge the shock created by Stuxnet, malware used to disrupt Iranian nuclear program in 2010. Industrial Control Systems are frequently called Cyber-Physical Systems (CPS) because they consist of software and network components deeply embedded in the physical world. Examples of such systems include water treatment, electricity generation/distribution, manufacturing, (petro)chemical production and other processes. Correspondently, attacks on CPS are called cyber-physical attacks. This talk presents the evolution of the ICS exploits and tactics to picture the ongoing "Race-to-the-Bottom" trend between ICS threat actors and defenders. There are two conclusions which follow this talk: (1) Traditional IT security approaches are not enough to defend against cyber-physical attacks and defenses should additionally include process- and control-engineering methods, (2) Due to the potential of cyber-physical attacks to have kinetic effect and cause casualties, it is urgent and of utmost importance for the international community of IT security specialists, governments and humanitarian lawyers to have a conversation about how to regulate the deployment of cyber-physical attacks.

Introduction

In the Information Technology (IT) domain, increasingly there is a gap between the attacker and defender capabilities. The attackers embraced firmware modifications and supply chain rootkits a decade ago while the defense community has recently embraced data diodes and application whitelisting. Current IT security defense technologies are not matched to offensive capabilities of threat actors and the gap keeps increasing, slowly

becoming hard to close. It is highly likely that a similar pattern will repeat for industrial control systems, and it is hoped that understanding the historical trends of the IT security industry will provide a discussion point when anticipating threats and planning defense strategies for Industrial Control Systems against cyber-physical attacks.

History of IT security

Security is a moving target. At first security was introduced into the network to prevent hackers from stealing passwords and impersonating communication parties or Man-in-the-Middle (MITM) attacks. Later the security moved into the computer and the operating system (OS). As the attackers became practiced in exploiting OS, security controls had to expand into software applications, resulting in such solutions as sandboxing and hypervisors. However, even these technologies are no longer enough.

The fundamental flaw in modern defensive computer security is the assumption that a personal computer (PC) is only a single computer running a single operating system. The fact is that nearly every hardware component that used to be "dumb" has been replaced with a "smart" component. For example, network cards now have their own firmware (own OS), built-in web server and perform complex cryptographic tasks. The main CPU of a computer system requests these other "computers" for access and data. A modern computer is not just one computer anymore!

The advantage for the attacker is that these other computers lack almost all the security protections built into modern operating systems. It is currently less labor intensive to write and maintain a rootkit for a firmware than to maintain the same router kiting functionality in the main operating system.

In the hacking community it is sometimes called *Race-to-the-Bottom*. As soon as security is introduced at some layer of computer or network architecture abstraction, the attackers are placing their exploits one layer down. While Windows got its own firewall not long ago, the attackers are already mastering their skills in exploiting silicon microchips.

Current trends in ICS security

Industrial control system (ICS) is a collective term used to describe different types of control systems and associated hardware instrumentation,

software applications and communication infrastructure, which are used to automate and operate industrial processes. Some of the industrial processes are called critical infrastructures because they are critical to the wellbeing of the population (e.g. water and power distribution utilities).

Industrial automation has followed the IT hardware development path. What used to be a simple analog sensor is now an IP-enabled smart transmitter with complex firmware, multiple wired and wireless communication modes, a large number of configuration possibilities, and even a web-server so that maintenance staff could calibrate and configure the device without approaching it.

The sensors used in ICS are also becoming more distributed than ever before with new types of sensors being introduced continuously. Tank farms are frequently placed in safer locations away from the main production plants. Weather sensors are placed outside the plant fence. Predictive maintenance systems with additional sensors are being integrated into assets that were previously only mechanical machines. We should take a look at the current trends in ICS exploitation to see whether industrial controller and smart field instrumentation (smart sensors and actuators) could become an attacker target any time soon.

The major difference between IT and cyber-physical attacks is the attacker's end goal. While in the IT domain the ultimate goal of the attacker is to get access to certain data, in the ICS domain the attacker's goal is to cause impact in the physical world.

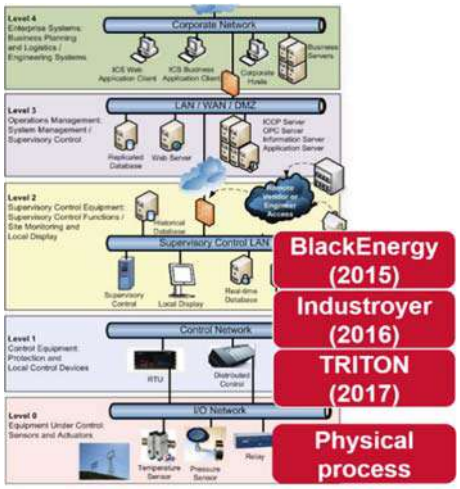
Besides Stuxnet (www.langner.com/wpcontent/uploads/2017/03/tokilla centrifuge.pdf), there were three other cyber-physical attacks in the past years:

- Attacks on three power substations in Ukraine, 2015; malware family BlackEnergy3;[1]
- Attack on power substation in Ukraine, 2016; malware family – Industroyer;[2]
- Attack on Safety Instrumented System (SIS) in a Saudi Arabia refinery, 2017; malware family TRITON.[3]

[1] https://ics.sans.org/media/EISAC_SANS_Ukraine_DUC_5.pdf.

[2] www.welivesecurity.com/wp-content/uploads/2017/06/Win32_Industroyer.pdf.

[3] www.fireeye.com/blog/threatresearch/2017/12/attackersdeploynewicsattackframework-triton.html.



If we map these attacks against the Purdue reference model of the Industrial Control Systems network architecture, we will notice that with each attack threat actors are moving their exploits one network layer lower. Thus, the first attack on Ukrainian power grid in 2015 (BlackEnergy3) was executed at the level of Human Machine Interface (HMI) by taking control over the operator's screen. In the second attack (Industroyer), threat actors moved one layer lower and launched their exploits at the level of industrial control protocols. In the TRITON attack, threat actors attempted to inject malicious code directly into the memory of Programmable Logic Controller (PLC) belonging to Safety Instrumented System (SIS), thus placing themselves very closely to the I/O cards and field instrumentation (sensors and actuators). It means, that the attackers have moved their exploits to tTRITON attackhe immediate proximity of the physical processes. This is because majority of embedded systems (controllers and field instruments) are currently lacking any exploit mitigation capabilities and defenders have little experience in performing compromise assessment and forensic analysis on these systems.

To date, it is unclear how to evaluate these attacks from the legal perspective. During the attack on Ukrainian power grid in 2015, around 250k people were left without power supply for an hour. At that time no national government condemned the execution of this attack and as a result, attacks on critical infrastructures were silently accepted as a "new normal" (this phenomenon is sometimes called "normalization of deviance"). A similar attack on the Ukrainian power grid in 2016 affected 225k people and was similarly left undiscussed from the legal standpoint. The situation became more critical in 2017, when the attackers targeted Safety

Instrumented Systems with TRITON exploit. SIS are designed to guard civilians in hazardous facilities from e.g. toxic releases or explosions and are meant to prevent casualties. By targeting these systems, the threat actors are willingly putting human lives in danger and denying their right to be safe. This is why it is of utmost importance to regulate deployment of cyber-physical attacks by the means of international laws.

Miscellaneous

Cyber-physical attacks are not the only attacks which can put human lives in danger. On June 27th, 2017, a massive cyberattack (NotPetya ransomware) hit Ukraine on the eve of Constitution Day. The attack payload was distributed via the updates of the most popular tax-filling software in Ukraine. Once delivered to the organization's network, the malware spread rapidly across all reachable computer systems. Cash counters in the grocery stores, bank ATMs, medical facilities, metro ticket sales points, critical infrastructures, industrial production enterprises, governmental organizations – all these systems were paralyzed by the malware. The life in the country almost entirely came to holt in less than 24 hours after the time of attack initiation.

> In Ukraine, the country's Minister of Health, Ulana Suprun, told BBC Future that her office was taken about 30 years back in time. "We're working by pen and paper again," she says. "There are so many things we can't do because we're down," says Suprun.
> For example, her Ministry centralizes the distribution of medicine across the vast territory of Ukraine's 24 regions. When hospitals in those regions run low on medications for patients, they contact the Ministry to source medicine. Either the Ministry has them, or they locate them in other regions and send them to the region in need.
> "But we can't relay those messages right now, except by phone, so imagine how crippling that is to us," says an exasperated Suprun. "What used to require one email, copied to the 24 regions now requires 24 separate phone calls before we can find the drugs. Ukrainians can't get medical documents because our internal system is down. I can't pull up statistics for a meeting I have this week about Aids. I couldn't even tell you which hospitals went down because they can't reach us".[4]

Through the globally interconnected systems, NotPetya rapidly propagated worldwide, causing significant downtime and financial losses to

[4] www.bbc.com/future/article/20170704-the-day-a-mysterious-cyber-attack-crippled-ukraine.

such multinational companies as Maersk, pharmaceutical giant Merck, FedEx's European subsidiary TNT Express, French construction company Saint-Gobain, food producer Mondelēz, manufacturer Reckitt Benckiser and many others. Currently, the world has very little understanding of such international IT infrastructure interrelationships and dependencies. Maersk has suffered 300 million in recovery costs due to the NotPetya attack.[5]

It is critical to understand that at the time of national or global crisis the speed of the infrastructure recovery will be dependent on the number of available professionals to perform these tasks. Typically, companies rely on external service providers for recovery operations. At the time of crisis when multiple organizations are affected, the amount of trained work force will become a major bottleneck. Additionally, the majority of organizations have weak backup policies and processes, which may result in unavailability of backups. This, in turn, will lead to significant delays in recovery and return to operational state.

Exploitation of embedded systems (industrial equipment, IoT and mobile devices, automotive systems and others) once used to be an exotic skill. These days, cyber-criminals and state-sponsored threat actors demonstrate advanced skills in exploiting these systems. An example includes the VPN Filter attack (2018), in which more than 500,000 SOHO (small and home office) routers of diverse makes/models worldwide were targeted by sophisticated modular malware with a multistage payload.[6]

Moreover, as the attackers are becoming more skilled, it is getting easier for them to exploit large organizations such as vendor or service providers to get access to a small number of intended targets. Recently, for example, attackers subverted the ASUS software update process to distribute their malicious code. It turned out, that the adversaries were merely interested in about 600 specific MAC addresses of ASUS laptops or in another words – in only 600 intended targets (www.wired.com/story/asus-software-update-hack/). In general, usage of staging targets such as equipment vendors or trusted service providers became a preferable method of the attackers to get access to their intended targets. In 2018 USA and UK governments have issued security advisories to warn industrial control systems and critical infrastructure operators about state-sponsored attacks with the goal of intrusion and establishing persistence in the infrastructure, possibly for future needs.[7]

---

[5] www.reuters.com/article/maersk-results-idUSL8N1L21HM.

[6] http://blog.talosintelligence.com/2018/05/VPNFilter.html.

[7] www.us-cert.gov/ncas/alerts/TA18-074A and www.ncsc.gov.uk/news/joint-us-uk-statement-malicious-cyber-activity-carried-out-russian-government.

Conclusions

While the world has not yet discovered a disruptive attack code hidden in the smart field instrumentation, it does not mean that threat actors are not capable of exploiting these systems or fitting complex exploitation codes into resource-constrained devices. Thus, the number of attacks on various embedded systems has exploded in the past years. Also, researchers have already shown that it is possible to embed the entire code for sophisticated physical damage attack into the firmware of a smart transmitter.

Following IT domain trends, the ICS defenders mount firewalls and harden the systems to prevent the attackers from reaching lower layers of industrial control networks. In response, the attackers are searching for alternative exotic pathways into the control network, where the supply chain, the external maintenance laptop and subcontractors' remote access are just a few examples. Overall, supply chain compromise is becoming one of the most serious threats to industrial environments with limited opportunities for proactive detection and prevention of these attacks.

The origin of one of the TRITON attacks was narrowed down to Central Scientific Research Institute of Chemistry and Mechanics in Moscow, Russia. This is a previously unseen *modus operandi* for offensive operations, when an intrusion team was moved closer to the team of technologists (engineers) and indicates a formation of multidisciplinary teams when conducting cyber-physical attacks. Note, that other national states have previously claimed their capability to disrupt civilian and critical infrastructures by the means of cyberattacks at the time of political or military crisis.

Defending against a sophisticated attacker require sophisticated defense methods. In addition to canonical IT security protections, it is important to include engineering approaches from the process- and control-engineering domain. For example, forged sensor readings can be detected via plausibility and consistency checks – the same methods, which are used for detecting faulty sensors. Predictive maintenance algorithms could be used for spotting early signs of process or equipment degradation so that operators could take corrective actions in a timely manner and prevent an attacker from completing their destructive mission.

Due to the potential of cyber-physical attacks to have kinetic effect and cause casualties, it is urgent and of utmost importance for the international community of IT security specialists, governments and humanitarian lawyers to have a conversation about how to regulate the deployment of cyber-physical attacks to prevent potential humanitarian crises.

Utilisation contemporaine et future des technologies cyber/numériques dans les conflits armés

Camille FAURE
Directrice adjointe, Direction des affaires juridiques,
Ministère des Armées français

Introduction

Mesdames, messieurs, chers amis,

Je suis très honorée de pouvoir intervenir aujourd'hui dans le cadre de cette table ronde. Je tiens à remercier tout particulièrement le Professeur Fausto Pocar pour son invitation, ainsi que tous les organisateurs de cette manifestation.

Le thème de la place de l'homme et des défis posés au droit international humanitaire par l'utilisation de nouvelles technologies dans les conflits retenu cette année fait écho à plusieurs travaux de fond conduits ou en cours au sein du ministère des armées sur les conditions dans lesquelles selon la France, le droit international s'applique aux opérations cyber ou numériques en temps de paix ou en temps de conflit armé, sur les opérations spatiales militaires et ou encore sur l'intelligence artificielle.

En préparant cette intervention, je me suis demandée si l'Institut de San Remo n'avait pas été victime d'un *malware* qui avait ingénieusement accentué la difficulté du sujet qu'il me revient de traiter, celui de l'« *Utilisation contemporaine et future des technologies cyber/numériques dans les conflits armés* ».

En effet, bien des distinctions ou des définitions du sujet sont peu aisées à cerner.

La distinction entre les usages contemporains et futurs des technologies numériques bute sur un triple obstacle :

i) celui du temps tout d'abord. En matière de technologies, créées le plus souvent par le secteur civil, la distinction entre aujourd'hui et demain est totalement poreuse. Demain est déjà hier. Les différences d'estimation avancées par les experts sont considérables : on évalue par exemple la croissance d'ici 2020 du nombre d'objets connectés (qui constituent une source de vulnérabilité en raison des interconnexions qu'ils permettent) à une fourchette située entre 26 et

212 milliards. Que certains experts dans ces augustes assemblées veuillent bien m'excuser si selon leur avis, ce que j'identifie comme un usage futur est selon eux déjà contemporain voire dépassé.

ii) Celui des usages ou utilisations ensuite :
- les usages sont tout sauf clairs et connus de tous. L'opacité en matière d'utilisation des technologies numériques, des dommages causés, des effets, est complète ;
- les usages sont parfois combinés avec les moyens cinétiques classiques, ce qui complexifie tant l'analyse que la prospective ;
- les usages par les états responsables et les acteurs non étatiques sont évidemment encore moins connus, analysés, décryptés, comme peuvent l'être les armes cinétiques.

Je note qu'il n'est question dans mon sujet que de technologies et pas d'armes.

Je ne doute pas que mon voisin (L. Gisel) évoquera la question difficile de la frontière entre l'utilisation de moyens numériques et d'une arme cyber qui intègre une technologie numérique. Il semble que l'on doive en tout état de cause se préparer à une généralisation de l'emploi demain des moyens numériques, dans les armes ou non, sur le champ de bataille virtuel ou réel, j'y reviendrai.

iii) Celui du conflit armé enfin.

Si je reprends les catégories juridiques bien connues de tous :
- s'agissant du *jus ad bellum* : à ma connaissance aucun Etat ne s'est officiellement déclaré victime d'une agression armée au sens de la Charte des Nations Unies du fait du recours à des moyens numériques, ou de moyens numériques combinés avec des moyens cinétiques atteignant un certain seuil de gravité (même si nous savons tous ici que l'interprétation juridique de ce terme fait l'objet de divergences de longue date entre Etats) ;
- s'agissant du *jus in bello* : cœur de nos débats, le manque de transparence sur ce point est complet, à la fois sur les utilisations actuelles, sur les recherches sur des armes futures et sur les stratégies des acteurs étatiques et non étatiques en ce domaine ;
- cependant, nous savons tous que la plupart des cas d'usage présumés de ces technologies se situent dans une zone grise qui se loge entre la criminalité économique ou financière, la manipulation de l'information ou la pénétration non autorisée. Tous ces usages sont situés en deçà de la menace ou du recours à la force au sens de l'article 2.4 de la Charte des Nations Unies. Ce qui pose la question cruciale des contre-mesures nécessaires et

proportionnées que les Etat pourraient envisager, en sus de la mise en cause de la responsabilité internationale de l'Etat auteur, et sur lesquels ils ne communiquent pas.

Ces questionnements permettent d'identifier trois grandes catégories d'enjeux associés aux utilisations contemporaines et futures des technologies numériques dans les conflits armés qui relèvent :

- de questions capacitaires pour les Etats, et plus généralement interrogent les missions westphaliennes de ces derniers ;
- de l'anticipation par les Etats de nouvelles formes de conflits, pour réduire leur risque d'occurrence et réfléchir sur le cadre de la conduite des hostilités ;
- de la régulation enfin, de tous les acteurs : crime organisé, groupes armés non étatiques, *proxies,* Etats, sociétés commerciales afin de tendre à une sécurité globale dans un contexte de vulnérabilité croissante.

Il me semble que la réponse à la question des usages actuels et futurs des technologies numériques dans les conflits armés peut être placée sous le pavillon d'un double paradoxe :

- celui d'une part du lien entre progrès technique croissant, vulnérabilité et efficacité de l'emploi des technologies numériques dans les conflits armés,
- celui d'autre part de l'absence de renouvellement des fondamentaux des conflits armés par les scénarios d'un emploi généralisé et massif des technologies numériques. Aussi, l'enjeu est moins l'emploi et les techniques d'aujourd'hui et de demain, qui seront certes fulgurants, que la question des seuils et de la prévisibilité des réactions à ces emplois, pour éviter un embrasement par accident.

Le progrès technique constitue le premier facteur de développement des usages des technologies numériques dans les conflits, dont l'efficacité est décuplée par la vulnérabilité des acteurs publics et privés

La technologie numérique : rupture technologique et rupture d'emploi

Une rupture technologique caractérisée notamment par 4 points :

- la sophistication croissante des moyens d'agression qui exige une grande maturité technologique, pour assurer une maîtrise et un contrôle stricts de l'emploi de moyens numériques
- leur forte évolutivité et leur expansion, facteurs de nivellement technologique et d'attrition de la supériorité militaire étatique *:* la nature de la menace, les outils, les tactiques, les techniques et procédures évoluent rapidement et se complexifient *;*
- l'immédiateté et les effets multiples de l'action – y compris à distance - dans un milieu opaque *:* les effets sont d'ordre physique – neutralisation d'un système d'arme – ou immatériel –collecte de renseignement –, temporaires, réversibles ou définitifs ;
- des difficultés d'attribution.

Trois ruptures d'emploi sont favorisées notamment par la dissémination de cette technologie :

- un emploi offensif et défensif, direct ou indirect (via des partenaires ou des industriels) contre des cibles très larges, de moyens dont l'efficacité est décuplée
- un jeu complexe des acteurs *:*
  Les acteurs étatiques *:* contribuent directement à ces ruptures technologiques et d'emploi en diffusant des armes cybernétiques qui, peuvent être réutilisées. Ils laissent transiter ou se développer sur leurs territoires les moyens de ces attaques.
  Ces acteurs poursuivent les trois grands objectifs opérationnels offensifs, classiques (renseigner, défendre, agir) et défensifs pour protéger leurs intérêts et réseaux.
  Des proxys, allant de diasporas instrumentalisées à des milices ou groupes armés, capables de tenir en échec des forces classiques, qui bénéficient de la dissémination des armes par les Etats eux-mêmes.
- une asymétrie et une désinhibition croissante dans l'emploi de la violence, au mépris du droit des conflits armés *:* un acteur disposant de capacités offensives, mais qui offrirait une surface de vulnérabilité numérique moins étendue, peut s'engager à moindre risque dans une escalade conflictuelle contre un Etat par exemple

Le numérique constitue une source de vulnérabilité systémique qui encourage le recours à ces moyens dans le cadre de conflits armés

La vulnérabilité systémique se caractérise par : l'insuffisance de l'état de sécurité, la numérisation massive des données, l'inter-connectivité croissante des réseaux, les fonctionnalités superflues de produits

polyvalents, des dizaines de milliards d'objets connectés en 2020 (26 à 212mds), des attaques facilitées via l'internet des objets.

La vulnérabilité publique et militaire est une réalité en raison du progrès technique.

Le numérique constitue le substrat sans lequel aucune activité ou mission ne peut plus être correctement conduite.

Deux exemples :

- le combat collaboratif (programme SCORPION) permet la mise en réseau des actions de combat, le partage de la connaissance de la situation sur un terrain entre des plates-formes facilitant une prise de décision rapide, l'intégration entre les plates-formes terrestres, les drones et les hélicoptères de combat, au sein d'une armée ou entre armées alliées ;
- le combat aérien du futur reposera sur le « *fonctionnement collaboratif en réseau, la guerre électronique, de radar, d'optronique de connectivité et de discrétion* » (cf. papier défense et innovation, ministère des armées, page 10).

La perspective de l'emploi généralisé et massif des technologies numériques ne renouvelle cependant pas les fondamentaux des conflits armés. L'enjeu est moins l'emploi et les techniques que la question des seuils et la prévisibilité des réactions en rappelant le droit international applicable pour éviter un embrasement par accident

Scénarios pour un futur emploi : le développement d'armes de « désorganisation massive »

Intérêt de la recherche universitaire pour ce sujet. Je vous renvoie aux travaux du « *center for long term cyber security* » de l'université de Berkeley dont un document publié en février 2019 détaille 4 scénarios en matière de cyber sécurité à l'horizon 2025 :

- Prolifération des techniques : l'accès à la technologie des ordinateurs quantiques et leur plus ou moins grande prolifération au profit d'acteurs criminels ou étatiques (capacité à casser les logiciels de cryptage, articulation avec l'Intelligence Artificielle).
- Cyber technologies comme source de conflit dans un monde dominé par les ordinateurs et l'intelligence artificielle, qui met un terme à toute incertitude et à tout aléa induit par l'action humaine. Effet induit en termes d'accroissement des conflits territoriaux (plus d'ambiguïté, évaluation parfaite des dommages à l'environnement, *etc.*).

- Le questionnement des acteurs : l'inefficacité des Etats et acteurs privés pour réguler et assurer la sécurité des systèmes ; le jeu des acteurs non étatiques et des organisations criminelles ; suscite deux réponses possibles, les Etats ne s'investissent plus et délèguent aux grands acteurs privés la sécurité (cf. la mise en œuvre 1996 manifeste de John Perry Barlow, "*A declaration of independence in Cyberspace*" ; au motif que cela constitue un risque en termes de prévisibilité des réactions), ou le maintien d'un pouvoir sur le cyber et sa régulation est perçu comme l'un des attributs essentiels d'un Etat souverain (approche westphalienne).
- L'éclatement d'internet : régulation de la sécurité par la seule intelligence artificielle, émergence de plusieurs internet plus ou moins sécurisés.

Ces scénarios sont certes un peu connexes à notre sujet mais intéressants car ils posent trois questions : celle des acteurs, de l'avenir de l'Etat westphalien, des vulnérabilités. Dont j'ai parlé plus haut.

S'agissant des documents publics d'Etats sur leurs scénarios concernant les conflits armés numériques de demain, je vous renvoie au Livre blanc de 2013, à la Revue stratégique de 2017, à la Revue cyber de 2018, qui dessinent le panorama suivant sur la vision française :

- Le livre blanc de 2013 envisage des attaques massives, constitutives d'actes de guerre.
- La revue stratégique s'inscrit dans la même ligne : « *Dans le cyberespace, certaines attaques, en raison de leur ampleur et de leur gravité, pourraient relever de la qualification d'agression armée : une attaque informatique majeure, par les dommages qu'elle causerait, pourrait ainsi justifier l'invocation de la légitime défense au sens de l'article 51 de la Charte des Nations Unies* ».
  Elle précise les technologies et précision des armes dont l'emploi est prévisible :
  - des ordinateurs quantiques permettant de casser les algorithmes de cryptographie actuels
  - des technologies de rupture et des innovations civiles
  - la combinaison de l'Intelligence artificielle et du cyber. « *Et songez à la combinaison future d'attaques cyber et d'intelligence artificielle, se livrant à un combat sur les réseaux à une vitesse défiant toute compréhension humaine* »

- des évolutions de doctrine du combat : le combattant serait recentré sur des missions de haute valeur ajoutée, moins dangereuses.

La France a créé récemment une « *Red team* », cellule de 4 à 5 personnes capables de proposer des scénarios de disruption (échafauder des hypothèses stratégiques valides sur des technologies disruptives de nature à bouleverser les plans capacitaires, réfléchir aux usages asymétriques de ces technologies) afin d'orienter les efforts d'innovation.

Cependant, les usages futurs s'inscrivent dans une zone grise qui, comme sur les théâtres classiques, ne renouvelle pas les fondamentaux des conflits armés

Il appartient aux Etats de s'adapter à cette zone grise qui emporte des enjeux capacitaires, organisationnels et opérationnels.

1) Des enjeux capacitaires : capacité de réponse à des menaces multiples qui se perfectionnent.
   La capacité de se protéger contre les attaques informatiques (réseaux des infrastructures critiques souveraines protégées par une posture permanente de cyber sécurité en France), de les détecter, d'en identifier les auteurs, de produire en toute autonome des dispositifs de sécurité (en matière de cryptologie et de détection d'attaque) est devenue un des éléments de la souveraineté nationale.
   Pour y parvenir, l'État doit soutenir des compétences scientifiques et technologiques performantes.
   La capacité d'attribution relève d'un besoin d'appréciation autonome, alors que les acteurs de ces théâtres se prêtent particulièrement à la clandestinité et à la manipulation.
   La capacité d'anticipation d'un déni de service, déni spatial partiel, total, d'origine physique ou cyber, C4 est cardinale.
2) Des enjeux d'organisation et opérationnels.
   L'établissement d'un commandement responsable au sens du droit international humanitaire est un enjeu de premier ordre. La France a par exemple fait le choix d'une organisation intégrée, unifiée, pour garantir une vision globale d'entrée et une mobilisation rapide des moyens nécessaires.
   La revue stratégique de cyberdéfense, publiée en février 2018, a élaboré une stratégie à part entière dans ce domaine en renforçant l'organisation de la cyberdéfense autour d'un centre de coordination des crises cyber et de quatre chaînes opérationnelles distinctes. En

complément des chaînes « protection », « renseignement » et « investigation judiciaire », la chaîne « action militaire » a recours à la lutte informatique offensive, y compris en soutien des actions de lutte informatique défensive (LID).

La France a ainsi consolidé un nouveau modèle de cyberdéfense, dont la création du commandement de la cyberdéfense (COMCYBER) en mai 2017 avait constitué une des étapes fondatrices.

L'enjeux spécifique des utilisations contemporaines et futures des technologies numériques dans les conflits armés réside dans la prévisibilité, dans les qualifications et seuils et dans le respect du DIH pour éviter un embrasement par l'effet d'une erreur d'évaluation

Il me semble, sans empiéter sur les propos de mon voisin, que la prévisibilité et la sécurité dans le domaine des technologies numériques, suppose qu'un maximum d'acteurs puisse répondre publiquement à trois grandes questions, ou donner des indications sûres :

- la catégorisation de l'échelle des menaces,
- la définition des seuils et l'application du droit international qui en résulte.

Une opération cyber peut être considérée comme un recours à la force prohibé au titre de l'article 2.4 de la Charte des Nations Unies. Le franchissement du seuil de l'emploi de la force n'est pas fonction du moyen cyber employé, mais des effets de la cyber opération. Si ces derniers sont similaires à ceux qui résultent d'armes classiques, l'opération cyber peut constituer un recours à la force. Dans le cyberespace, comme dans les autres domaines, le droit international existant s'applique et doit être respecté.

La France considère qu'une attaque informatique majeure, perpétrée par un Etat ou des acteurs non-étatiques agissant sous le contrôle ou les instructions d'un Etat , eu égard aux graves dommages qu'elle causerait (exemples : pertes humaines substantielles, dommages physiques considérables, déficience des infrastructures critiques avec des conséquences significatives), pourrait constituer une « agression armée », au sens de l'article 51 de la Charte des Nations Unies, et justifier ainsi l'invocation de la légitime défense.

Cette légitime défense peut être mise en œuvre par des moyens conventionnels ou cybernétiques pour peu que soient respectés les principes de nécessité et de proportionnalité. La caractérisation d'une attaque informatique en tant qu'« agression armée », au sens de l'article 51 de la

Charte des Nations Unies, relève d'une décision politique au cas par cas à la lumière des critères établis par le droit international.

A l'heure actuelle, les opérations de lutte informatique offensive sont concourantes aux opérations militaires conventionnelles. L'hypothèse d'un conflit armé constitué exclusivement d'activités numériques ne peut être exclue par principe, mais repose sur la capacité des opérations cyber à atteindre le seuil de violence requis pour qualifier l'existence d'un conflit armé international ou non-international.

Malgré leur caractère dématérialisé, ces opérations restent soumises au champ d'application géographique du droit international humanitaire, c'est-à-dire que leurs effets sont limités au territoire des Etats parties dans le cadre d'un conflit armé international ou sur le territoire sur lequel se déroulent les hostilités en conflit armé non-international.

Pour conclure, la France considère que l'émergence d'un cadre de cyber sécurité collective, ne pourra reposer que sur les équilibres définis par le droit international. La « Stratégie internationale de la France pour le numérique » souligne en outre l'importance pour la France de poursuivre « *un dialogue coopératif avec l'ensemble des acteurs privés et publics concernés, et l'ensemble des partenaires internationaux qui y sont prêts, sur le plan bilatéral comme multilatéral* ».

La France a pris une part active aux négociations conduites dans le cadre des cinq derniers groupes d'experts gouvernementaux sur la cyber sécurité au sein de l'ONU et en fera de même pour les nouveaux cycles qui s'engagent (groupe d'experts gouvernementaux et *open-ended working group*) cet automne. Elle est également engagée dans d'autres enceintes internationales où sont abordées ces questions de sécurité de l'espace numérique.

Afin de garantir un cyberespace ouvert, sûr, stable, accessible et pacifique, la France réaffirme son attachement à l'applicabilité du droit international, dont la Charte des Nations Unies dans son intégralité, du droit international humanitaire, de la Déclaration universelle des droits de l'homme et du droit international coutumier, à l'usage des technologies de l'information et de la communication (TIC) par les États.

A cet égard, je vous renvoie au papier sur le droit international appliqué aux opérations dans le cyberespace que les autorités françaises viennent de rendre public.

Merci de votre attention.

The use of cyber technology in warfare: which rules does IHL provide and are they sufficient?*

Laurent GISEL
Senior Legal Advisor, International Committee of the Red Cross

As part of the mandate it has been entrusted by the international community, the ICRC closely follows the development of new means and methods of warfare and their use by parties to armed conflicts. Its assessment of the foreseeable humanitarian impact of new means and methods of warfare, and the challenges they may pose to international humanitarian law (IHL), focuses on interrelated legal, military, technical, ethical and humanitarian considerations. The assessment of whether IHL rules are adequate and sufficient to protect civilians against the effects of cyber operations during armed conflicts is based on the potential human cost imposed by such operations, the protection that existing law affords, and the challenges these operations pose for the interpretation and application of IHL.

After a few introductory remarks on the reality of the use of cyber operations during armed conflict today, this presentation will, therefore, be divided in three parts. The first part will analyse cyber operations as such, in particular their potential human cost, and some of the technical characteristics that may raise concern. The second part will underscore the protection that existing IHL affords against the effects of cyber operations during armed conflicts. And the last part will address some of the challenges that cyber operations pose for the interpretation and application of IHL, to inform the discussion on whether the protection afforded by IHL may be deemed adequate and sufficient.

The use of cyber operations during armed conflicts

Most cyber attacks and other cyber operations that are reported around the world belong to the realm of cyber criminality, espionage or other type

* The views expressed in this presentation are those of the author alone and do not necessarily reflect the views of the ICRC. The author would like to thank Kubo Mačák and Tilman Rodenhäuser for their useful comments on this paper.

of cyber intrusions or incidents, and do not appear to be part of an armed conflict.[1]

However, the use of cyber operations during armed conflicts is also a reality. The U.S., the U.K. and Australia have declared using cyber operations during armed conflicts, in particular, against the Islamic State group.[2] Cyber operations have been used by Israel and the Hamas.[3] Cyber operations also affected other countries involved in armed conflicts, such as Georgia in 2008,[4] Ukraine in 2015-17[5] or Saudi Arabia in 2017.[6] However, the authors of these cyber attacks remain unknown and attribution of responsibility is contested. It is, therefore, not confirmed that these operations had a nexus to these armed conflicts. Finally, there have been media and other reports of cyber operations by States against other States,

---

[1] Gary P. Corn, 'Cyber National Security: Navigating Gray Zone Challenges In and Through Cyberspace', in Winston S. Williams and Christopher M. Ford (ed.s), *Complex Battlespaces: The Law of Armed Conflict and the Dynamics of Modern Warfare*, Oxford, Oxford University Press, 2018, p. 347; Kubo Mačák, 'From the Vanishing Point Back to the Core: The Impact of the Development of the Cyber Law of War on General International Law', in H. Rõigas et al (ed.s), *Defending the Core*, Tallinn, NATO CCD COE, 2017, p. 141.

[2] Mike Burgess, 'Offensive cyber and the people who do it', Australian signals Directorate, 27th March 2019, www.asd.gov.au/speeches/20190327-lowy-institute-offensive-cyber-operations.htm; Jeremy Fleming, 'Director's speech at CyberUK18', U.K. GCHQ, 12th April 2018, www.gchq.gov.uk/speech/director-cyber-uk-speech-2018; 'Statement of General Paul M. Nakasone, Commander, United States Cyber Command, before the Senate Committee on Armed Services', 14th February 2019 www.armed-services.senate.gov/imo/media/doc/Nakasone_02-14-19.pdf.

[3] BBC, 'Hackers Interrupt Israeli Eurovision WebCast with Faked Explosions', 15th May 2019, www.bbc.co.uk/news/technology-48280902; Zak Doffman, 'Israel Responds to Cyber attack with an Air Strike on Cyber attackers in World First', Forbes Magazine, 6th May 2019, www.forbes.com/sites/zakdoffman/2019/05/06/israeli-military-strikes-and-destroys-hamas-cyber-hq-in-world-first/#1c692f73afb5. While the purported target of the Hamas cyber operation has not been publicly released, the targeting of the Hamas Cyber HQ building by kinetic means was based on intelligence gained as part of the IDF cyber defence effort.

[4] David Hollis, 'Cyberwar Case Study: Georgia 2008', *Small War Journal*, 2010, https://smallwarsjournal.com/blog/journal/docs-temp/639-hollis.pdf.

[5] Andy Greenberg, 'How an Entire Nation Became Russia's Test Lab for Cyberwar', *Wired*, 20th June 2017, www.wired.com/story/russian-hackers-attack-ukraine/; Andy Greenberg, 'The Untold Story of NotPetya, the most devastating cyberattack in history', *Wired*, 22nd August 2018, www.wired.com/story/notpetya-cyberattack-ukraine-russia-code-crashed-the-world/.

[6] Blake Johnson, Dan Caban, Marina Krotofil, Dan Scali, Nathan Brubaker and Christopher Glyer, 'Attackers Deploy New ICS Attack Framework "TRITON" and Cause Operational Disruption to Critical Infrastructure', *Fireeye Blogs*, 14th December 2017, www.fireeye.com/blog/threat-research/2017/12/attackers-deploy-new-ics-attack-framework-triton.html.

in what is sometimes referred to as a 'grey zone'.[7] The analysis of this so-called 'grey zone' and the extent to which IHL may apply to cyber operations between States in the absence of on-going kinetic hostilities is outside the scope of this presentation.[8]

Beyond these declared or alleged uses of cyber operations during armed conflicts, an increasing number of States are developing military cyber capabilities, whether for offensive of defensive purposes. This increase in cyber capabilities also increases the likelihood that they will be used in future conflicts. So, let us first turn to the potential human cost of cyber operations.

The potential human cost of cyber operations

As noted in a recent ICRC position paper,[9] cyber operations may offer less destructive alternatives to other means or methods of warfare, but they also carry considerable risks. They may enable militaries to achieve their objectives without harming civilians or causing physical damage to civilian infrastructure. On the other hand, recent cyber operations show that sophisticated actors have developed the capability to negatively affect the provision of essential services to the civilian population.

To develop a realistic assessment of the potential human cost of cyber operations, in November 2018 the ICRC invited cyber security experts and cyber threat analysts from all parts of the world to share their knowledge about the technical possibilities, expected use, and potential effects of cyber

---

[7] See Camilla Faure, "Utilisation contemporaine et future des technologies cyber/numériques dans les conflits armés", above pp. 80-88; Gary Corn, "Punching on the Edges of the Grey Zone: Iranian Cyber Threats and States Cyber Responses", *Just Security*, 11 February 2020, www.justsecurity.org/68622/punching-on-the-edges-of-the-grey-zone-iranian-cyber-threats-and-state-cyber-responses/; Lindsey R. Sheppard, 'Warning for the Gray Zone', Center for Strategic and International Studies, Report, 13 August 2019, www.csis.org/analysis/warning-gray-zone.

[8] For the ICRC view on the matter, see the 2016 ICRC Commentary on Article 2 of the First 1949 Geneva Convention, para. 253-256. See also Michael N. Schmitt and Liis Vihul (ed.s), *Tallinn Manual 2.0 on International law applicable to cyber operations*, 2nd ed., Cambridge University Press, Cambridge, 2017 (Tallinn Manual 2.0), Rule 82; and 'Scenario 13: Cyber operations as a trigger of the law of armed conflict' in Kubo Mačák, Tomáš Minárik and Taťána Jančárková (ed.s), *Cyber Law Toolkit* (2019-), https://cyberlaw.ccdcoe.org/.

[9] ICRC, *International Humanitarian Law and Cyber Operations during Armed Conflicts*, Position paper, November 2019, p. 3, www.icrc.org/en/download/file/108983/icrc_ihl-and-cyber-operations-during-armed-conflicts.pdf (hereinafter: ICRC position paper).

operations. The report of the meeting, "*The potential human cost of cyber operations*" is available online.[10]

Cyber operations can pose a threat for critical civilian infrastructure. One area of concern for the ICRC, given its mandate, is the health-care sector. Research shows that this sector appears to be particularly vulnerable to cyber attacks, because of the increased digitization and interconnectivity in health care. Hospital medical devices are connected to the hospital network. Biomedical devices such as pacemakers and insulin pumps are sometimes remotely connected through the internet. This increases the sector's digital dependency and attack surface and leaves it exposed. The consequences of cyber attacks against medical facilities can quickly become significant, possibly life-threatening.[11]

Other types of critical civilian infrastructure, including electricity, water and sanitation, have also been affected by cyber attacks. The frequency of these attacks is reportedly increasing, and the severity of the threat evolved more rapidly than anticipated just a few years ago.[12]

Beyond the vulnerability of specific sectors, our research highlighted three technical characteristics of cyber operations that are particularly concerning.[13]

First, cyber operations carry a risk of over-reaction and escalation. It may be almost impossible for the target of a cyber attack to detect whether the attacker's aim is to spy or to cause physical damage. This might only be identified once a harmful effect on the targeted system is achieved. There is a risk that the target will anticipate the worst-case scenario and react much more strongly than it would have done if it had known that the attacker's true intent was limited to espionage, for example.

Second, cyber tools and methods can proliferate in a unique manner that is difficult to control. Today, sophisticated cyber attacks are only carried

---

[10] ICRC, *The potential human cost of cyber operations*, Expert Meeting Report (Laurent Gisel and Lukasz Olejnik, eds), May 2019, www.icrc.org/en/download/file/97346/the-potential-human-cost-of-cyber-operations.pdf (hereinafter: The potential human cost of cyber operations).

[11] Ibid, pp. 18-22.

[12] See Marina Krotofil, 'Casualties caused through computer network attacks: the potential human costs of cyber warfare' above pp. 73-80; see also: The potential human cost of cyber operations, note 10 above, pp. 23-28; Sergio Caltagirone, 'Industrial cyber attacks: a humanitarian crisis in the making', *Humanitarian Law and Policy blog*, 3 December 2019, https://blogs.icrc.org/law-and-policy/2019/12/03/industrial-cyber-attacks-crisis/.

[13] The potential human cost of cyber operations, note 10 above, p. 7.

out by the most advanced and best-resourced actors. But once a cyber tool is used, stolen, leaked or otherwise becomes available, other actors may be able to find it, reverse-engineer it and repurpose it for their own – possibly malicious – ends.

Third, attribution is challenging. It requires time, resources and expertise. Identifying actors who violate IHL in cyberspace and holding them responsible is likely to remain a challenge. While the principles and rules governing attribution of conduct for the purpose of State responsibility apply whether the conduct is carried out by cyber or any other means, the perception that it may be easier to deny responsibility for cyber attacks may make actors less scrupulous about respecting international law.[14]

To sum up, while cyber operations have caused and continue to cause massive economic costs, they have fortunately not caused major human harm so far. But they have exposed the vulnerability of essential services, and some of their technical features raise concern. Furthermore, as Durham has noted,[15] much is unknown in terms of technological evolution, the capabilities and the tools developed by the most sophisticated actors, and the extent to which the future, possibly more frequent, use of cyber operations during armed conflicts might qualitatively differ from the trends observed so far. Caution is therefore warranted not to draw hasty conclusions.

Having considered the potential human cost and certain technical characteristics of cyber operations, let us turn to the law. Two points will be addressed: the protection that IHL affords and whether IHL may be deemed adequate and sufficient.

The protections that IHL affords against the harmful effects of cyber operations

As noted above, the vulnerability of the health care sector to cyber attacks is a matter of concern. Happily, IHL provides crucial safeguards that go some way towards addressing this concern. For example, belligerents must respect and protect medical facilities and personnel at all

[14] ICRC position paper, note 9 above, pp. 8-9.
[15] Helen Durham, 'Opening Remarks', above pp. 18-22.

times.[16] This means that cyber attacks against the health care sector in an armed conflict would in most cases violate existing IHL. Likewise, IHL prohibits attacking, destroying, removing or rendering useless objects indispensable to the survival of the civilian population.[17]

More generally, IHL prohibits directing cyber attacks against civilians and civilian objects,[18] including civilian infrastructure. It prohibits indiscriminate and disproportionate cyber attacks.[19] It requires taking all feasible precautions in attack, in particular, in the choice of means and methods of warfare, to avoid or at least minimize incidental civilian harm.[20] The fact that cyber operations may offer alternatives to militaries to achieve their objectives without harming civilians or causing permanent physical damage to civilian infrastructure will become increasingly relevant over time with regard to this obligation to take all feasible precautions. Indeed, when the use of a cyber operation during an armed conflict helps to avoid or minimize incidental civilian harm compared to other available means or methods of warfare, and provided that such use is 'feasible' as understood

[16] See, for instance, Art. 19 of the 1949 Convention (I) for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field (GC I); Art. 12 of the 1949 Convention (II) for the Amelioration of the Condition of Wounded, Sick and Shipwrecked Members of Armed Forces at Sea (GC II); Art. 18 of the 1949 Convention (IV) relative to the Protection of Civilian Persons in Time of War (GC IV); Art. 12 of the 1977 Protocol Additional to the Geneva Conventions of 12th August 1949, and relating to the Protection of Victims of International Armed Conflicts (AP I); Art. 11 of the 1977 Protocol Additional to the Geneva Conventions of 12th August 1949, and relating to the Protection of Victims of Non-International Armed Conflicts (AP II); Henckaerts and Doswald-Beck (ed.s), Customary International Humanitarian Law, Vol. I: Rules, ICRC, Cambridge University Press, Cambridge, 2005, https://ihl-databases.icrc.org/customary-ihl/eng/docs/v1_rul (hereinafter ICRC Customary IHL Study) Rules 25, 28 and 29.

[17] Art. 54 AP I; Art. 14 AP II; Rule 54 ICRC Customary IHL Study.

[18] Art.s 48, 51 and 52 AP I; Rules 1 and 7 ICRC Customary IHL Study.

[19] Art. 51(4 and 5) AP I; Rules 11, 12 and 14 ICRC Customary IHL Study. Indiscriminate attacks are those: (a) which are not directed at a specific military objective; (b) which employ a method or means of combat which cannot be directed at a specific military objective; or (c) which employ a method or means of combat the effects of which cannot be limited as required by international humanitarian law; and consequently, in each such case, are of a nature to strike military objectives and civilians or civilian objects without distinction. Disproportionate attacks are those which may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects or a combination thereof, which would be excessive in relation to the concrete and direct military advantage anticipated.

[20] Art. 57 AP I; Rules 15 - 21 ICRC Customary IHL Study.

under IHL[21] in the circumstances ruling at the time, it becomes required as a matter of law.[22]

Furthermore, parties to the conflict must take all feasible measures to protect civilians and civilian objects under their control against the effects of hostilities.[23] This could include, for example, segregating military and civilian networks in the same way that military airports are separated from civilian commercial airports.[24]

Depending on how key IHL notions are interpreted for cyberspace (see below), the principles of distinction, proportionality and precautions have

---

[21] According to various declarations given by States when ratifying AP I or in military manuals, and to the definitions given in Protocols II and III to the 1980 Convention prohibiting Certain Conventional Weapons, feasible precautions are "those precautions which are practicable or practically possible taking into account all circumstances ruling at the time, including humanitarian and military considerations"; Protocol on Prohibitions or Restrictions on the Use of Mines, Booby-Traps and Other Devices (Protocol II to the 1980 CCW Convention), 1980, Art. 3(4); Protocol on Prohibitions or Restrictions on the Use of Incendiary Weapons (Protocol III to the 1980 CCW Convention), 1980, Art. 1(5); Protocol on Prohibitions or Restrictions on the Use of Mines, Booby-Traps and Other Devices as amended on 3rd May 1996 (amended Protocol II to the 1980 CCW Convention), Art. 3(10). For State practice, see the ICRC Customary IHL Study, p. 54.

[22] Art. 57, and, in particular, 57(2)(a)(ii) AP I. See Marco Roscini, *Cyber Operations and the Use of Force in International Law*, Oxford University Press, Oxford, 2014, p. 235. The U.S. Department of Defense 2015 Law of War Manual (up-dated December 2016; hereinafter U.S. DoD LoW Manual https://ogc.osd.mil/images/law_war_manual_december_16.pdf) discusses the use of cyber tools as potential measures to reduce the risk of harm to civilians or civilian objects, noting that "cyber capabilities may in some circumstances be preferable, as a matter of policy, to kinetic weapons" (para. 16.5.3.1). For this author, as noted above, it becomes a matter of law and not policy when the choice of using cyber capabilities is 'feasible' as understood under IHL. This debate has been accurately summarized by an International Law Association Study Group report with regard to another new technology, precision-guided munitions: "it bears emphasis that all precautionary obligations are 'technology neutral', i.e. they apply irrespective of the weapons technology used, including PGMs [Precision-Guided Munitions]. Thus, if the use of a PGM avoids or minimizes incidental civilian casualties compared to another means or method of warfare, and provided its use is feasible under the given circumstances (i.e. taking into account both military and humanitarian considerations), then using such a PGM is compulsory. Similarly, if the only way to carry out an operation without violating the prohibitions of indiscriminate or disproportionate attacks is to use a PGM, then the attacker is faced with only two options: to use the PGM; or not carry out the attack at all." (International Law Association Study Group, *The Conduct of Hostilities and International Humanitarian Law Challenges of 21st Century Warfare*, final report, 25th June 2017, p. 45, https://ila.vettoreweb.com/Storage/Download.aspx?DbStorageId=3763&StorageFileGuid=11a3fc7e-d69e-4e5a-b9dd-1761da33c8ab).

[23] Art. 58 AP I; Rules 22 to 24 ICRC Customary IHL Study.

[24] ICRC, *International humanitarian law and the challenges of contemporary armed conflicts*, report, 2015, p. 43, www.icrc.org/en/download/file/15061/32ic-report-on-ihl-and-challenges-of-armed-conflicts.pdf (hereinafter ICRC 2015 IHL Challenges Report).

the potential to afford strong protection against cyber attacks directed at critical civilian infrastructure, or against incidental civilian harm – including indirect effects – that may be expected when carrying out a cyber attack against a military objective.

Despite the interconnectivity that characterizes cyberspace today, these principles can indeed be respected. While some of the known cyber tools were designed to self-propagate and indiscriminately affect widely-used computer systems, they do not do so by chance: the ability to self-propagate normally needs to be specifically included in the design of the tool.[25]

In fact, many of the cyber attacks reported in public sources appear to have been rather discriminate from a technical perspective. This does not mean that they were lawful or would have been lawful if carried out during a conflict; on the contrary, a number of the cyber attacks that have been reported in public sources would be prohibited by IHL if carried out as part of an armed conflict because they apparently targeted civilian objects. However, their technical characteristics show that cyber operations may well be precisely tailored and create effects on specific targets only, rendering them capable of being used in compliance with IHL. Ensuring that cyber operations affect only the targeted object may, however, be technically challenging and require careful planning in their design and use.[26]

So, let us turn finally to the challenges that cyber operations pose for the interpretation and application of IHL.

Challenges posed by cyber operations for the interpretation and application of IHL

The identification of the challenges that cyber operations pose for IHL, and how they can be answered, is necessary to inform the analysis of whether IHL rules are adequate and sufficient, or whether they require to be reaffirmed, clarified, strengthened or developed further. In particular, there are a number of critical debates on how IHL applies to cyber operations, including on the notion of attack under IHL and whether data is an object for the purposes of IHL. At the very least, these debates underscore the need to clarify how IHL applies.

---

[25] Lukasz Olejnik & Tilman Rodenhäuser, 'Malware: A selection of essential cyber notions and concepts', *Humanitarian Law and Policy blog*, 23rd May 2019, https://blogs.icrc.org/law-and-policy/2019/05/23/malware-essential-cyber-notions-concepts/.

[26] ICRC position paper, note 9 above, p. 5.

The notion of attack under IHL

As noted above, the principles of distinction, proportionality and precautions have the potential to afford strong protection to critical civilian infrastructure. However, the most detailed and onerous rules stemming from these principles apply only to attacks,[27] and not all cyber operations are attacks under IHL, which are defined as "acts of violence against the adversary, whether in offence or in defence".[28] There is, therefore, a need to clarify which cyber operations amount to an attack under IHL.

What defines an attack is not the means or methods being used, but the effects or consequences that these means and methods have.[29] On the basis of this understanding, the Tallinn Manual 2.0 proposed the following definition of a cyber attack: "A cyber attack is a cyber operation, whether offensive or defensive, that is reasonably expected to cause injury or death to persons or damage or destruction to objects".[30]

It is widely accepted that cyber operations expected to cause death, injury or physical damage constitute attacks under IHL. It is the prevailing view among the academic community. Only a very few official State documents such as military manuals have so far provided for a definition of cyber attack under IHL or offered examples of cyber operations that would qualify as attacks. All those that do provide a definition appear to encompass at least cyber operations causing such effects.[31] Some of

---

[27] See text in relations to note 18 to 20 above.

[28] Art. 49 AP I. The notion of attack under IHL, as defined in Art. 49 AP I, is different from and should not be confused with the notion of 'armed attack' under Art. 51 of the UN Charter, which belongs to the realm of *jus ad bellum*. To affirm that a specific cyber operation, or a type of cyber operation, amounts to an attack under IHL does not necessarily mean that it would qualify as an armed attack under the UN Charter, and many of them would not.

[29] Cordula Droege, 'Get off my cloud: cyber warfare, international humanitarian law, and the protection of civilians', 94 *International Review of the Red Cross*, 2012, p. 557; Danish Ministry of Defence, Defence Command Denmark, *Military Manual on International Law Relevant to Danish Armed Forces in International Operations*, 2016, pp. 290-291, https://fmn.dk/eng/allabout/Documents/Danish-Military-Manual-MoD-defence-2016.pdf (Danish Military Manual); Tallinn Manual 2.0, commentary on Rule 92, para. 3.

[30] Tallinn Manual 2.0, Rule 92.

[31] Australian Government, *Australia's International Cyber Engagement Strategy*, 2017, Annex A, p. 90, www.dfat.gov.au/sites/default/files/DFAT%20AICES_AccPDF.pdf; Danish Military Manual, pp. 290-291; France, *International Law Applied to Operations in Cyberspace*, 2019, p. 13, www.defense.gouv.fr/content/download/567648/9770527/file/international+law+applied+to+operations+in+cyberspace.pdf; Norway, *Manual i krigens folkerett*, 2013, para. 9.54 https://fhs.brage.unit.no/fhsxmlui/bitstream/handle/11250/194213/manual_krigens_folkerett.pdf?sequence=1&isAllowed=y (Norwegian Military Manual) New Zealand Defence Force, *Manual of Armed Forces Law, Volume 4, Law of Armed Conflict*, DM

them[32] expressly include harm due to the foreseeable indirect (or reverberating) effects of an attack, and the ICRC shares this view.[33] This could be the case for example of the death of patients in intensive-care units caused by a cyber operation against an electricity network that results in cutting off a hospital's electricity supply.

Beyond this, it is well known that there are different views among the experts who drafted the Tallinn Manual 2.0 and in other academic circles on whether a cyber operation that disables an object without physically damaging it amounts to an attack under IHL.[34] Because cyber operations can significantly disrupt essential services without necessarily causing physical damage, this constitutes one of the most critical debates for the protection of civilians against the effects of cyber operations.

The definitions adopted in the military manuals of Norway and New Zealand mirror the definition adopted by the Tallinn Manual 2.0. As the commentary on the relevant rule of the Tallinn Manual 2.0 precisely exposes the various views of how damage should be understood in this context, it is unclear whether these manuals were meant to express a position on this debate. Australia considers cyber operations as attacks if they rise "to the same threshold as that of a kinetic 'attack under IHL'",[35] but, again, it is unclear whether this was meant to take a position in this debate.

The Danish military manual specifies with regard to attacks that "[a]s far as damage to objects is concerned, the term covers any physical damage. However, the term does not cover temporary inoperability and other neutralization which does not involve physical damage (e.g., a digital "freeze" of a communication control system)."[36] The U.S. Department of Defense Law of War Manual does not propose a definition. Along similar lines, however, when discussing cyber operations that constitute attacks under IHL it provides as an example a "cyber attack that would destroy

69 (2nd ed.), Volume 4, 2017, para. 8.10.17, www.nzdf.mil.nz/downloads/pdf/public-docs/dm_69_2ed_vol_4.pdf (New Zealand Military Manual); US DoD LoW Manual, para. 16.5.1.

[32] Danish Military Manual, p. 677 (when discussing computer network attacks); New Zealand Military Manual, para. 8.10.22; Norwegian Military Manual, 9.54.

[33] ICRC position paper, note 9 above, p. 7.

[34] Tallinn Manual 2.0, commentary on Rule 92, para.s 10 to 12.

[35] Australian Government, *Australia's International Cyber Engagement Strategy*, 2017, Annex A, p. 91.

[36] Danish Military Manual, p. 290. The Manual specifies with regard to computer networks attacks and operations that "[t]his means, for instance, that network-based operations must be regarded as attacks under IHL if the consequence is that they cause physical damage" (p. 291; footnote removed).

enemy computer systems", and notes that "[f]actors that would suggest that a cyber operation is not an 'attack' include whether the operation causes only reversible effects or only temporary effects".[37] However, these two manuals do not clarify what they mean by "reversible" or "temporary", or whether – and if so when – a long-lasting effect may no longer be deemed temporary. In this respect, it is also important to note that the possibility to repair the physical damage caused by a military operation (whether cyber or kinetic) to an object is not generally understood as a criterion disqualifying the operation as an attack under IHL. This is the case even if the repair would reverse the direct effect of that operation and restore the functionality of the object in question.[38]

France has expressed a clearer and broader understanding of the notion of cyber attack. It "considers that a cyber operation is an attack where the targeted equipment or systems no longer provide the service for which they were implemented, whether temporarily or permanently, reversibly or not. If the effects are temporary and/or reversible, the attack is characterised where action by the adversary is necessary to restore the infrastructure or system (repair of equipment, replacement of a part, reinstallation of a network, etc.)."[39] Commenting this position, Schmitt noted that "[t]his view is highly defensible as a matter of law, for the plain meaning of damage reasonably extends to systems that do not operate as intended and require some form of repair to regain functionality".[40]

For the ICRC, to consider a cyber operation as an 'attack' only if it causes bodily harm to humans or physical damage is a too much of a restrictive understanding. It ignores the harmful effects that can be caused through cyber operations without causing physical damage. Such a narrow

---

[37] U.S. DoD LoW Manual, para.s 16.5.1 and 16.5.2 respectively.

[38] For example, Michael Lewis discusses the practice of conducting bridge attacks longitudinally during the 1991 Gulf War, and, among other, notes that "damage to the bridge would be nearer midspan and therefore more easily repaired", without claiming that this quality would prevent the operation to qualify as an attack ('The Law of Aerial Bombardment in the 1991 Gulf War', *American Journal of International Law*, Vol. 97, No. 3 (2003), pp. 481–509, at 501). See also Michael N. Schmitt, 'France Speaks Out on IHL and Cyber Operations: Part II', *EJIL:Talk!*, 1 October 2019, www.ejiltalk.org/france-speaks-out-on-ihl-and-cyber-operations-part-ii/ who argues that 'the plain meaning of damage reasonably extends to systems that do not operate as intended and require some form of repair to regain functionality'.

[39] France, *International Law Applied to Operations in Cyberspace,* 2019, p. 13.

[40] Michael N. Schmitt, 'France Speaks Out on IHL and Cyber Operations: Part II', *EJIL:Talk!*, 1 October 2019, www.ejiltalk.org/france-speaks-out-on-ihl-and-cyber-operations-part-ii/.

understanding would be difficult to reconcile with the object and purpose of the rules on the conduct of hostilities, which is to ensure the protection of the civilian population and civilian objects against the effects of hostilities.[41] If the electricity supply to the civilian population is cut, it will indeed affect the population in the concerned area in a similar manner until electricity is restored, regardless of whether the electricity cut was caused by the use of a traditional bomb, a graphite bomb, or a cyber operation (though the means or method used might impact on the speed, means and expertise needed to regain functionality and restore the supply).

To sum up, it is the ICRC's position that key IHL rules may only provide the full scope of legal protection if States recognize that cyber operations that impair the functionality of objects, such as civilian critical infrastructure, are subject to the rules governing attacks under IHL.[42]

The question of how widely or narrowly the notion of 'attack' is interpreted with regard to cyber operations is, therefore, essential for the protection that IHL rules afford to civilians and civilian infrastructure. It would thus seem critical for the international community to work towards building consensus on the interpretation of this notion. This will impact the assessment of whether or not existing rules are adequate and sufficient to protect civilians against the effects of cyber operations considering their specific technical characteristics. This also raises the question of the rules governing cyber operations other than attacks.

Cyber operations other than attacks

To identify, and possibly clarify, the rules that govern cyber operations other than attacks is an issue requiring more attention than it has received so far. This is all the more critical if only those operations that cause physical damage would be understood as attacks: in this case, the category of cyber operations other than attacks would be significantly broader and could cause more serious negative effects for the civilians and civilian infrastructure.

The various specific protection regimes under IHL often afford protection that goes beyond the protection against attacks, including certain very important safeguards. This is the case of the protection afforded to the medical mission, to objects indispensable to the survival of the

[41] ICRC position paper, note 9 above, p. 7-8.

[42] ICRC, *International humanitarian law and the challenges of contemporary armed conflicts*, report, 2019, p. 28, www.icrc.org/en/document/icrc-report-ihl-and-challenges-contemporary-armed-conflicts.

population,[43] or to the delivery of humanitarian assistance,[44] among others. But what about the application of the principles of distinction, proportionality and precaution to operations other than attacks?

Some military manuals express the view that cyber operations other than attacks may be directed against civilians or civilian objects.[45] At least for States parties to the First Additional Protocol, this would seem difficult to reconcile with the 'basic rule' expressed in Article 48, which provides that "parties to the conflict… shall direct their operations only against military objectives".

Beyond this basic rule, it is interesting to observe that references are made to the principle of military necessity. For instance, Australia notes that "[a]pplicable IHL rules will also apply to cyber operations in an armed conflict that do not constitute or rise to the level of an 'attack', including the principle of military necessity (…)".[46] The U.S. DoD Law of War Manual specifies that "such operations [cyber operations that do not amount to an attack under IHL] must not be directed against enemy civilians or civilian objects unless the operations are militarily necessary".[47] While these references to military necessity as a restraining principle are welcome, more clarity is needed on exactly what the principle of military necessity entails when conducting cyber operations.

Furthermore, IHL requires that constant care must be taken to spare the civilian population, civilians and civilian objects in the conduct of military operations.[48] This obligation extends to military operations other than attacks.[49] As explained in the UK military manual, this means that the

---

[43] See, among others, text in relation to notes 16 and 17 above.

[44] See Art.s 23 and 59 GC IV, 69 and 70 AP I; Rules 31 and 32 ICRC Customary IHL Study; Rule 145, Tallinn Manual 2.0, which refers to cyber operations, and not only cyber attacks ("Cyber operations shall not be designed or conducted to interfere unduly with impartial efforts to provide humanitarian assistance."; see also commentary para. 4 on Rule 80: "Certain cyber operations, such as those affecting the delivery of humanitarian assistance (Rule 145), are governed by the law of armed conflict even if they do not rise to the level of an 'attack'").

[45] Norwegian Military Manual, para. 9.57; US DoD LoW Manual, para. 16.5.2.

[46] Australian Government, *Supplement to Australia's Position on the Application of International Law to State Conduct in Cyberspace*, 2019, p. 4, www.dfat.gov.au/sites/default/files/application-of-international-law-to-cyberspace.pdf.

[47] US DoD LoW Manual, para. 16.5.2.

[48] Art. 57(1) AP I; Rule 15 ICRC Customary IHL Study; Rule 114, Tallinn Manual 2.0.

[49] Tallinn Manual 2.0, commentary on rule 114, para. 2. The ICRC Commentary on Art. 57 AP I, para. 2191 defines 'military operations as follows: "The term 'military operations' should be understood to mean any movements, manoeuvres and other activities whatsoever carried out by the armed forces with a view to combat."

commander must "bear in mind the effect on the civilian population of what he is planning to do and take steps to reduce that effect as much as possible."[50] The view expressed in the U.S. DoD Law of War Manual is also worth noting, according to which a cyber operation that is not an "attack" "should not be conducted in a way that unnecessarily causes inconvenience to civilians or neutral persons".[51]

It is clear from these few elements that cyber operations other than attacks are not unregulated. However, the legal regime governing such operations remains less complete than the legal regime governing operations that amount to attacks under IHL.

Qualification of data as an object under IHL

Another issue that requires clarification is the protection that IHL affords to data. Some of the specific protections afforded by IHL extend to essential data, such as data belonging to medical units, which are in the ICRC's view encompassed in the obligation to respect and protect such units.[52] The same can be said of data necessary to the delivery of humanitarian relief.[53]

However, many forms of data that are necessary for the functioning of society and the provision of essential services to the civilian population do not necessarily enjoy specific protection under IHL. Today, social security data, tax records, civilians bank accounts or companies' client files are all digitalized. Deleting or tampering with these data could quickly bring government services and private businesses to a complete standstill. They could cause more harm to civilians than the destruction of physical objects.

---

[50] UK, *The Joint Service Manual of the Law of Armed Conflict*, JSP 383, 2004, para. 5.32.1.

[51] U.S. DoD LoW Manual, para. 16.5.2; see also Schmitt's proposal that "States would commit, as a matter of policy, to refraining from conducting cyber operations to which the IHL rules governing attacks do not apply when the expected concrete negative effects on individual civilians or the civilian population are excessive relative to the concrete benefit related to the conflict that is anticipated to be gained through the operation" (Michael N. Schmitt, "Wired warfare 3.0: Protecting the civilian population during cyber operations", *International Review of the Red Cross* (2019) 101 (1), 333–355, p. 347, https://international-review.icrc.org/sites/default/files/reviews-pdf/2019-12/irrc_101_910_17.pdf).

[52] ICRC 2015 IHL Challenges Report, note 24 above, p. 43.

[53] See Tilman Rodenhäuser, "Hacking Humanitarians? IHL and the protection of humanitarian organizations against cyber operations," *EJIL:Talk!*, 16 March 2020, www.ejiltalk.org/hacking-humanitarians-ihl-and-the-protection-of-humanitarian-organizations-against-cyber-operations/.

Assuming they do not fulfil the criteria for being a military objective, why would such data be less protected than their physical equivalent?[54]

The few views expressed so far by States diverge. The Danish military manual considers that "[g]enerally speaking, however, (digital) data do not in general constitute an object".[55] Conversely, the Norwegian military manual holds that data shall be regarded as objects and may only be attacked directly if they qualify as a lawful target.[56] France takes what could be seen as a middle view when stating that "[g]iven the current state of digital dependence, content data (such as civilian, bank or medical data, etc.) are protected under the principle of distinction."[57]

Here again, it would be critical to work towards building consensus among the international community on how IHL protects civilian data in an increasingly digitalised world.

Avenues to reduce the potential human cost of cyber operations

Working towards the clarification of key IHL notions is critical, but various other measures may also be taken to reduce the potential human cost of cyber operations. Our May 2019 report listed many proposals and suggestions in terms of possible avenues to avoid or reduce the human cost of cyber operations,[58] and it is outside the scope of this presentation to discuss them all. Two of them will be mentioned as examples only.

First, already in peacetime, States may take a variety of measures in terms of precautions against the effects of attacks.[59] This could be the case, for example, of the creation of a "digital watermark" to identify certain actors or infrastructure in cyberspace, such as objects that enjoy specific

---

[54] Helen Durham, 'Opening Remarks', above pp. 18-22; ICRC position paper, note 9 above, p. 8. See also Kubo Mačák, 'Military Objectives 2.0: The Case for Interpreting Computer Data as Objects under International Humanitarian Law', *Israel Law Review*, Vol. 48, No. 1, 2015, pp. 55-80, at 77-80 (arguing that by considering data as an object, civilian data remains protected from attack and from excessive incidental harm in accordance with IHL's central value of protection of civilians and civilian objects); 'Scenario 12: Cyber operations against computer data' in Kubo Mačák, Tomáš Minárik and Taťána Jančárková (ed.s), *Cyber Law Toolkit* (2019-), https://cyberlaw.ccdcoe.org/.

[55] Danish Military Manual, p. 292.

[56] Norwegian Military Manual, § 9.58.

[57] France, *International Law Applied to Operations in Cyberspace,* 2019, p. 14.

[58] The potential human cost of cyber operations, note 10 above, pp. 75-77.

[59] ICRC 2015 IHL Challenges Report, note 24 above, p. 43.

protection under existing IHL.[60] Other possible measures include segregating military from civilian cyber infrastructure and networks; segregating from the internet computer systems on which essential civilian infrastructure depends; and more generally working on the cyber resilience of essential infrastructure. Many States and private businesses are developing standards and best practices in this regard, to respond to all kinds of cyber threats, including peacetime ones.[61]

The second example is the issue of addressing the unique way cyber tools proliferate, in particular, because of the ability to repurpose them. The actors who develop or use malware can make repurposing more difficult, for example, through encrypting the payload. This could prevent at least some actors from repurposing the tools and reduce the risk of subsequent misuse that their proliferation entails.[62] In fact, militaries say they do analyse the risk that cyber tools be reverse-engineered before taking the decision to use them, which probably operates as a restraining factor regarding their use.

International law does not prohibit repurposing cyber tools, a technique also helpful in cyber security testing. The actor that reverse engineers, reengineers, repurposes or otherwise reuses a malware is responsible for such use, including when it constitutes a violation of IHL. There is currently no specific IHL rule that would impose a residual responsibility on the actor that originally developed or used the malware. Yet, the idea that a belligerent might continue to bear some responsibility after having used specific means of warfare is not foreign to IHL. This is, for example, the case for explosives remnants of war.[63] Nothing would prevent States from deciding to move in this direction with regard to cyber tools if they deemed it appropriate.[64] They might also put in place mandatory equity

[60] The potential human cost of cyber operations, note 10 above, pp 40-41.

[61] See for example the 2016 *Directive (EU) 2016/1148 of the European Parliament and of the Council of 6 July 2016 concerning measures for a high common level of security of network and information systems across the Union* (NIS Directive).

[62] The potential human cost of cyber operations, note 10 above, p. 42.

[63] See the Protocol on Explosive Remnants of War (Protocol V to the 1980 CCW Convention), 28 November 2003, https://ihldatabases.icrc.org/applic/ihl/ihl.nsf/INTRO/610.

[64] Laurent Gisel & Lukasz Olejnik, 'Potential human costs of cyber operations—Key ICRC takeaways from discussion with tech experts', *Humanitarian Law and Policy blog*, 29th May 2019, https://blogs.icrc.org/law-and-policy/2019/05/29/potential-human-costs-cyber-operations-key-icrc-takeaways-discussion-tech-experts/.

processes to determine whether and when a particular vulnerability would have to be disclosed to the relevant software developer.[65]

Conclusion

A more in-depth discussion is needed, in particular, among States, on how IHL is to be interpreted for cyber operations during armed conflicts.[66] In this regard, the ICRC welcomes the renewed efforts of the international community to discuss how international law, including IHL, applies to cyber operations, whether through the Open-Ended Working Group and the Group of Government Experts created by the United Nations General Assembly,[67] or within other international organisations. For example, already a few years ago, the Asian-African Legal Consultative Organization created an Open-Ended Working Group on International Law in Cyberspace and has been regularly discussing the issue since then;[68] in 2018, the Commonwealth committed to move forward the discussion on how international law, including IHL, applies to cyber operations;[69] and the OAS's Inter-American Juridical Committee is leading a project focused on

[65] See The potential human cost of cyber operations, note 10 above, pp. 9, 33–34. See also Norm (j) of the norms, rules and principles for responsible behaviour of States proposed by the Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security in its 2015 report (UN Doc. A/70/174, para. 13(j) p. 8) and welcomed by the UN General Assembly in its Resolution A/RES/73/27, OP 1, pt 1.11 "States should encourage responsible reporting of ICT vulnerabilities and share associated information on available remedies for such vulnerabilities to limit and possibly eliminate potential threats to ICTs and ICT-dependent infrastructure."; and the norm for States to create a vulnerability equities process put forward by the Global Commission on the Stability of Cyberspace, *Advancing Cyberstability,* Final report, November 2019, pp. 38-39, https://cyberstability.org/wp-content/uploads/2020/02/GCSC-Advancing-Cyberstability.pdf ("States should create procedurally transparent frameworks to assess whether and when to disclose not publicly known vulnerabilities or flaws they are aware of in information systems and technologies. The default presumption should be in favor of disclosure.")

[66] ICRC position paper, note 9 above, p. 9.

[67] See United Nations General Assembly Resolutions A/RES/73/27 and A/RES/73/266.

[68] See e.g. Asian-African Legal Coordination Organization, *International Law in Cyberspace,* AALCO/58/ DAR ES SALAAM /2019/ SD/17, www.aalco.int/Final%20 Cyberspace%202019.pdf.

[69] Commonwealth Cyber Declaration, *Issued at the Commonwealth Heads of Government Meeting, London, United Kingdom, 16 – 20 April 2018*, https://thecommonwealth.org/commonwealth-cyber-declaration.

improving the transparency of how States understand international law in cyberspace.[70]

Today, in intergovernmental discussions a lot of attention is given to the question of whether IHL applies to cyber operations during armed conflict. Affirming that IHL applies in cyberspace is, however, only a first step. Clarifying how key IHL notions apply in this space is necessary to assess the extent to which IHL restricts military cyber operations, and whether IHL may be deemed adequate and sufficient. This in turn will inform the analysis of whether new rules might be useful or even needed. However, if new rules are developed, they should build upon and strengthen existing law. The ICRC stands ready to lend its expertise to such discussions.

[70] See OAS, 'International Law and State Cyber Operations: Improving Transparency (presented by Dr Duncan B. Hollis)', OAS Doc. CJI/doc 570/18, 9th August 2018, www.oas.org/en/sla/iajc/docs/CJI_doc_570-18.pdf.

III. IHL and new technology. How much human control is required by existing rules?

IN THIS SESSION, EXPERTS CONSIDERED A SCENARIO INVOLVING THE USE OF AUTONOMOUS WEAPON SYSTEMS AND SHARED THEIR VIEWS ON WHAT KIND OF HUMAN CONTROL IS NECESSARY TO ENSURE COMPLIANCE WITH IHL AND ETHICAL ACCEPTABILITY[*]

Chair:
Netta GOUSSAC
Legal Advisor, Arms Unit of the Legal Division, ICRC

Argument that IHL requires significant human control over weapon systems and decisions on the use of force

Richard MOYES
Managing Director, Article 36

Argument that IHL does not require significant human control over weapon systems and decisions on the use of force

Michael MEIER
Colonel (retd), Special Assistant for Law of War Matters, Office of the Judge Advocate General, International and Operational Law Division, US Department of the Army, Member, IIHL

Netta GOUSSAC

This session is about autonomous weapons systems. It's really my pleasure to be at the Sanremo Round Table for the very first time. I want to thank the Institute for including the topic of autonomous weapons on this year's agenda. I gather it's not the first time that the issue has been discussed here, but I think it's an important one to be included, particularly in light of this year's theme. As Helen Durham said yesterday, the

[*] The following discussion, based on the transcript of the recorded session, reflects the debate among the panelists. It has not been revised by them and does not commit them with regard to the views expressed.

development and use of autonomous weapon systems - weapons that are capable of selecting and attacking targets without human intervention - both spark people's imagination but also stir up existential questions and fears. On the one hand, autonomy in targeting functions of weapons may entail a number of advantages from a military perspective. But on the other hand, experts including the ICRC have raised concerns about a growing risk that humans will lose control over weapon systems and eventually become so far removed from the choice to use force that life and death decisions will effectively be left to sensors and software.

I also think that including this topic on this year's agenda is timely. In the context of inter-governmental discussions that have continued for six years now, notably under the auspices of the Convention on Certain Conventional Weapons, States have agreed on the importance of human judgment, responsibility, accountability when it comes to the use of autonomous weapon systems. The central question is now, what type and how much control are required over weapon systems and over the use of force in order to ensure compliance with international humanitarian law? And importantly for ethical acceptability. This is what we're going to examine in this session. In this session we will use practical scenarios that involve the use of autonomous weapons in order to examine what kind of practical measures can and should be taken to ensure compliance with IHL and ethical considerations.

It's an honor for me to moderate a session among two experts whose work I admire very much, Richard Moyes and Michael Meier. Richard, sitting to my right, is the Co-founder and Managing Director of Article 36, a UK-base Article 36, relevantly for this session, is also part of the leadership group of the campaign to stop killer robots. Under Richard's leadership, Article 36 developed the concept of "meaningful human control" as an approach to concerns regarding autonomy in weapon systems – and this is something we will examine closely today – and in fact Richard has played a leading role in several recent IHL-related developments including the Safe Schools Declaration, the Treaty on the Prohibition of Nuclear Weapons, and the Convention on Cluster Munitions.

Michael Meier, sitting to my left, is a member of this Institute as well as a senior civilian advisor to the US Army Judge Advocate General on matters related to the law of armed conflict. Prior to taking up his current role in the Department of the Army, he served more than two decades as an Army Judge Advocate before joining the US State Department as Attorney Adviser for political military affairs. Of course, in this room we have the privilege to have access to so much knowledge, expertise and experiences.

We would really like to hear from you in this session as well and we look forward to your contributions towards the end of the session.

To start with, what we wanted to do is to be clear about what we're talking about when we speak of autonomous weapons systems. Indeed, the absence of a common language when talking about these weapons has been a real hindrance in inter-governmental discussions. So many of the terms that we have been using and that we will be using today have been under the microscope, unpacked and, even at times, contested. At this point I should say thank you very much to the interpreters for bearing with us as we try to be clear about the language. But to start with, I wanted to pose a question to Michael. What are autonomous weapons? And are we speaking here of a closed class of weapons systems?

Michael MEIER

I wanted to first take the opportunity to thank the Institute and the organizers for this invitation to speak here today. It's nice to be back in Sanremo once again. I also wanted to clarify a couple of points. As Netta pointed out, I am with the Department of the Army and if you look in your book it says Department of Defense, but it is Department of the Army. And second and most importantly, while at times I will discuss and try to explain the US Government, Department of Defense and Department of Army positions related to this topic, I am speaking in my personal capacity and my views do not necessarily reflect those of the US Government, the Department of Defense or the Department of the Army. Essentially, you can explain it this way. If I say something that the Department of Defense agrees with, they will be more than happy to accept it and go with it. If I say something that they don't like very much, like Ethan Hunt in Mission Impossible, I will be disavowed and they will not accept anything that I say.

With that, I'd like to start on our question. What are autonomous weapons? I think Netta pointed out that this has been going on for six years and there's still no agreed definition of what an autonomous weapon at this point is. At least within the United States and the Department of Defense, we have a DoD directive 3000.09 which does define autonomous weapons as "a weapon system that once activated can select and engage a target without further intervention by a human operator. This does include human supervised autonomous weapons systems that are designed to allow human operators to override the operation of the weapon system but can select an engaged target without further human input after activation." My three years when I was at the State Department, part of the three years, I was part

of the CCW delegation. I started with the delegation in 2014 when these discussions began and stayed with them till I left the State Department in 2016. The big piece was what is an autonomous weapon? I think, instead of trying to work out a definition which some people have viewed as an impediment, I think it has actually been more helpful because then you've ensured that you've talked about the key issue, which is ensuring that weapons systems help effectuate the intent that the commander and the operators of the system. It's what the United States refer to as "human machine teaming", "human machine interaction", taking practical steps to reduce the risk of unintended engagements. So, I think the lack of definition has enabled participants within CCW and others to get a better understanding of the complexity involved with this type of technology.

Netta GOUSSAC

Thanks, Michael. So, from what you've said, autonomy isn't necessarily the central word to understand when talking about autonomous weapon systems. I wanted to pose a question to Richard, what is autonomy? What is the difference with automation? And is this the yardstick that we should be looking at when talking about these systems?

Richard MOYES

I would certainly agree that the sort of terminology in this debate has been unstable and contested and it makes conversations more complex, not just because individual words don't have a fixed meaning, but because people come to the subject matter with quite different conceptions of what we should be talking about. Broadly speaking, I don't think it's necessarily very useful to think of a distinction between automatic and autonomous in this context. I feel that the sort of definition that Michael laid out is quite useful in the sense that for me, all the systems that are under consideration in this area, within their process of operation, use sensors. There's an internal analysis of sensor data within the system and as a result of that, under certain conditions, force is applied by the system at a certain sort of target.

I think that's the key technological structure and that's the starting point in this debate for me, namely, the sort of closed loop between sensor inputs, machine analysis, and an application of force.

Obviously, this creates quite a broad categorization. There's already quite a lot of military systems that function in this way, in a way, from landmines, to certain types of anti-armor system, to missile problematic defence systems. So, putting these things together under this broad

umbrella, isn't to say that all these things are necessarily fundamentally. However, I do think that when we look at international legal instruments on specific weapons, we find quite often that weapon systems, where there's a direct loop between sensors, machine analysis and the application of force, have quite often been subject to specific regulations because I think there are certain tensions inherent in that. But I think we need to start with that broad framing of the technological category and not get too focused on the more futuristic anxieties around machines that operate completely independently of human command.

Essentially, I think professor Dinstein in his comments yesterday noted that a certain type of general AI was not going to be present in the near future. I think for us in civil society our concern is more about a trajectory and a movement towards some sort of greater autonomy in weapon systems from where we are now and perhaps an erosion of human control in that process. It's more at that level that our concerns, I think, are relevant rather than an anxiety about highly futuristic systems, which I don't think are a pressing concern and which I also don't think we can approach conceptually without having dealt with this more basic category.

Netta GOUSSAC

So, bringing that all together, what we're talking about are systems that use sensors in order to select and engage or attack targets from within a predefined group of targets without human intervention after they've been activated. As Richard mentioned, the concern at least from civil society is about incremental change or increases in autonomy, to use that word. But in fact, this broad conception of autonomous weapons includes some systems that are already in use today and from their use we can learn a little bit about how autonomous weapon systems are or should be used. So, I wanted to ask Michael, what are some current applications of autonomous weapons systems that could maybe ground our discussions today? And what is the reason that autonomy is being pursued by militaries?

Michael MEIER

Certainly, I think one of the reasons that autonomy is being pursued by militaries is that artificial intelligence and autonomy and other time-related technologies use software machine control rather than manual control by a human being. And these technologies can produce greater accuracy, precision and speed in weapon systems. It also allows us to produce entirely new capabilities that otherwise would be impossible if you were relying solely on humans. For example, the counter rocket artillery and

mortar system referred to as C-RAM is able to fire precisely at incoming projectiles and disable them, something that a human operator cannot do manually. So, you have various types of systems out there. I don't want to go into too many of the different weapons systems, but I do agree. I think the point is we have systems now that have autonomous features in them. I think it has been a good basis to inform the discussion because again, if you're talking about a future capability, one way to understand it is understand that we have been using autonomy and technology for decades and doing it safely. This will allow us to figure out how these other systems will work with less human involvement at the tail end.

Netta GOUSSAC

So, from what you've said, autonomous weapons systems already exist, not necessarily being viewed as problematic, but there is a drive to pursue future technological developments. Richard, did you have any views on that drive?

Richard MOYES

I think we also recognized that there are certain militarily useful capabilities that come with technologies in this area. I think it's been actually very positive in the Convention on Conventional Weapons that certain States have come with examples. The US has certainly done that and provided a sort of structured analysis of how those systems are managed in their operation -I would say how they're controlled, but I know control is a contested word in this space- but those systems are managed in a way that allows the application of the law.

One thing I was just going to point out in terms of military utilities as well, that I think States draw upon, is a sense that these systems can be used to strike at targets where the actual location of that target is not known or where even the presence of such a target is not necessarily known for certain. It seems to me that one of the key dynamics of these types of system, one of the sources of tension in relation to these systems, is that they involve an application of force at a specific time and a specific location that is not set by the human commander, but that occurs within some sort of envelope of operation that the human commander has put into process. Now, that has a utility, but perhaps it also brings with it necessarily certain tensions when it comes to evaluations of the likely effects that are going to occur from the use of that system.

Netta GOUSSAC

I think that's a really good point to talk about, the legal and ethical perspectives on the use of autonomous weapon systems. Because, as Richard pointed out, one of the notable features of using these systems is the distance - sometimes in space, but importantly the distance in time - between the decision to use or activate the system and the eventual use of force against a target, if at all. Michael, does this pose a challenge for the interpretation and application of IHL?

Michael MEIER

I think the shortest answer would be no, but then we would have to stop and we'd all go to lunch. So, I think, going back to what Professor Dinstein said yesterday, going back to the fundamental principles of IHL, the principles of distinction, proportionality and precaution, they provide the fundamental framework that governs the use of autonomy in weapons systems. They're consistent with also a military doctrine of effective and safe combat operations, independent of legal considerations. Sound military doctrine condemns the use of indiscriminate and excessive force as costly, inefficient and a waste of resources.

So, I think that when you're developing these systems you can look at military necessity. What is the need that the system is going to fail? Distinction: the system would need to be able to distinguish between combatants and civilians. You would look at proportionality when you're developing this and prohibit excessive incidental incidents. You can look at humanity: this would reduce unnecessary suffering. And, of course, the principle of honor, which is ensuring respect for IHL.

I think when you're looking at autonomy and AI in weapon systems, and you want to reduce the risk to civilians, there are ways that it does that. As we've already discussed, you've had self-destruct, self-deactivation systems that reduce the risk to civilians. Increasing awareness is one way that autonomy can help do that. Nowadays, casualties are often caused from the lack of awareness of civilians on the battlefield or the fog of war. Commanders may not be aware that civilians are there and in good faith may misidentify civilians as combatants. I think what we're seeing now with artificial intelligence is that humans would have to search through large amounts of data. For example, the United States currently has 11,000 unmanned aerial systems in place. They accumulate hundreds of thousands of hours of video every year. To put it into perspective, I read in an article, they said that each day we bring enough high resolution of video that

equates to three seasons of American football. It would take 20 analysts, 24 hours to go through less than 10% of that.

So, that's what the Department of Defense is doing with Project Maven, using artificial intelligence to go through this large amount of data to allow for precautions and better battlefield awareness. Soldiers are already using these sensors that Richard talked about. Currently, an infantry squad uses 150 different satellites to operate. A brigade combat team uses over 2,500 satellite systems in its operations. So, technology is being used to give commanders and operators a better view of the battlefield and it can also improve assessments to comply with the law of war. We have the joint targeting process in place - the United States uses, NATO uses - where sophisticated algorithms and software are already used to determine what types of munitions should be used and the effect of those munitions. Automated targeting identification, tracking, selecting, engaging targets, again, can allow us to strike objectives more accurately. The AMRAAM incorporates an active radar in its inertial reference system, which allows it to use its missile to guide it to the proper target. So, using the technology in consistence with the law of armed conflict is the goal, I think, of all responsible militaries.

Netta GOUSSAC

I want to ask a follow-up question there, because many of the issues that you've raised now are about identifying and selecting targets in accordance with the rules of IHL. But, of course, the rules of IHL also apply to the decision to apply or use force against the target. How do autonomous weapons systems challenge the ability to comply with IHL with respect to that final stage of the decision to use force? Given that, as you pointed out, those decisions are incredibly context-specific - they require an understanding of the context - but they are also rather qualitative in that they cannot be completely boiled down to the kinds of objective indicators that might be identified by sensors or images.

Michael MEIER

Certainly, I think that engaging the target, like you said, is the critical point. But I think the misperception is that it's sort of a binary choice that you're going to have humans involved or not. In at least the United States' view, in our view it's not replacing humans. You're giving humans better control. The human commander is still going to determine when the system's going to be employed and when it's going to be engaged. So, the human still has the ability to direct how the system's going to be used.

There are going to be temporal types, there are going to be geographical types of restrictions that the commander will impose on the system to ensure that it's use is consistent with the law of armed conflict.

Netta GOUSSAC

So, we're going to turn to those constraints when we go through the practical scenarios in a moment. But before that, Richard, I wanted to ask you, what motivated you and Article 36 to get involved in this discussion? What were the challenges and risks that you saw?

Richard MOYES

Sure. We certainly wouldn't argue that there are no potential benefits from new technology in the battle space. And when I say benefits, I guess I primarily mean benefits for civilian protection although it may obviously work positively for militaries as well.

What we're mainly concerned with in a way - as we were suggesting at the beginning - are those systems where you have this closed loop between sensors analysis and the application of force. That's something that's happening for me downstream of a human decision to put a system at work in a sort of operational space, so somebody has undertaken an attack with the system. These systems I think really have two structural elements that are significant and that for me create tensions regarding a legal application. One of those systems has to have embedded within it some sort of model or representation of what it is that is going to be attacked. These are systems which are going to attack a thing or a phenomenon that is detected by the sensors and there has to be some encoding in the system of the conditions upon which force will actually be applied. Now, encoding objects in the world into a sort of embedded model within a system always creates some tension regarding what falls within the model and what falls outside the model.

Coming back to the most basic examples, we might think of an anti-vehicle mine as being designed to detonate upon contact with a vehicle. But of course, the anti-vehicle mine simply detonates when pressure from a certain weight bears upon the mine - if a weight above 50 kilos, for example, bears upon the mine, that triggers the application of force. So there's a sort of target profile within that structure which is weight-graded at 50 kilos - which is not exactly the same as an armored vehicle or a vehicle - and there's tension because machines, computers, machines in general, don't see things in the world the way that we see them, process them. They have to reduce them to some sort of sensor-identifiable

representation, that, of course, has to be in the language that the sensors of that system use: for a mine, it is detecting weight. It can't detect heat signature because it only detects weight. So, the sensors you have condition how you can model the world and construct representations of it. Sometimes things will fall within that model that you don't actually want to attack, sometimes things that you do want to attack will fall outside of that model and a commander needs to understand how that model works. Not just what the system is intended to do in a theoretical sense, but how that target profile relates not just to things they may want to attack, but also to things that they may not wish to apply force to if they're going to make an informed legal judgment. Understanding that and the comprehensibility of the conditions upon which force will be applied I think is a very significant issue. I also, at a sort of ethical level, have anxieties about reducing people to a configuration of data points that a sensor simply identifies and processing that without some further human moral engagement with the process. Processing that through a machine system, to an application of force, to a person doesn't fall to my mind as a concern deriving straightforwardly from the rules of IHL. But it's a concern that I think can be invoked in people due to a sense that this involves ultimately machines reducing people to data representations and applying force to them. Still, however, there's a commander who has put this at work, so there isn't a complete absence of human engagement in that.

I agree with Michael that the parameters of time and space, duration and geography over which the system operates are fundamentally important. In a way, the embedded encoding of the target profile is a set of assumptions that a commander is putting to work. That set of assumptions doesn't just apply universally over all the area of land and over all time. It's a set of assumptions that needs to be bounded to a certain location if reasonable legal evaluations are going to be undertaken.

I think Helen Durham mentioned yesterday the agreement that the law is generally addressed to humans and it's humans who have to apply the law. I think we all would agree on that. But these issues of geographical and spatial constraint cause for me some anxieties about the structure of the law in the future. Because if humans, either individually or institutionally, have to make decisions about the legality of attacks on an attack-by-attack basis under the law, or if there are rules that relate to an attack, there needs to be some conceptual geographic-spatial bounding to what should be considered as an attack.

If you have a machine system that can go on and apply force in various locations against various object types and you're going to treat the use of

that system as one attack, you're actually seeing a much greater range of applications of force under that sort of legal concept than there might be sufficiently reasonable for the human legal judgments to be sufficiently informed, and to be really aware of what the contextual factors are in the situations where force is actually going to be applied. Because in any of the systems, specifically where force is going to apply and specifically when force is going to be applied may be uncertain bounded within those spatial and geographical constraints. So, I feel that it should be encouraged for States to adopt a recognition that the geographical and the durational boundaries of the use of a system need to be controlled, such that the human commanders can meet their established legal obligations.

Netta GOUSSAC

Thanks Richard. So it seems to me that when we're speaking of IHL, we're not talking about autonomous weapon systems making proportionality assessments or making distinction assessments, but rather we're talking about the commander, the person who is deciding on the use of the autonomous weapon system and making an appropriate choice of means and methods of warfare, being able to limit the effects of the weapon that they're using in accordance with IHL and in particular making the legal judgments that are required of them by IHL, notably from the rules of distinction, proportionality and precautions in attack.

I think it's probably a good moment now to move this discussion from the theoretical to the somewhat practical by introducing two practical scenarios involving the use of autonomous weapon systems. We'll take each one in turn and while we're discussing them, I also encourage you to think through how you would approach these scenarios and maybe share those views with us and your questions when we open for questions immediately afterwards. So, the first scenario, I'll read it out for the purposes of translation. Scenario one: during the course of an international armed conflict State A seeks to gain control of a major road between the airport and a major city in State B. State A wants to attack State B's military vehicles on that road. State A has the option of using an aerial loitering weapon system that loiters over a defined area for a specified time period. The system is capable of selecting and attacking a target without human intervention at any point after launch within the defined area, once it identifies a target signature. The system uses image recognition to identify predefined types of military vehicles belonging to State B based on their shape as well as thermal sensors to detect vehicle heat signatures. While there is reasonable confidence that the system will identify tanks and

armored vehicles, some civilian SUVs and trucks have similar shape and heat signatures to military jeeps and trucks. Michael, I'll start with you, what factors in your view should weigh on State A's decision on whether or not it would be lawful to use the autonomous weapons system in this scenario? And if State A does choose to use the weapon, what is needed to ensure compliance with IHL?

Michael MEIER

I think it's important to realize there are various factors that are going to go into this decision. The key here again, will be the human-machine interaction to ensure that the weapons system is going to help effectuate the intention of the commander and the operators of the system. What that means is the system's going to do what you designed it and what you want it to do. It's important to look at the situation holistically. I think this is part of the reason the United States doesn't like the warm human control, but appropriate levels of human judgment and human judgment over the use of force, because the State didn't just decide to roll this system out on the spur of the moment. This has gone through years of research, development and testing before the system would ever be deployed. And many of the questions that are being asked here would hopefully have been addressed during that process. So, I think when you start looking at it, with respect to profiles and tasking, not all weapons systems are going to be appropriate for all circumstances. Going back to the targeting process and military decision-making processes, these planning processes would help you determine whether or not this is the right system for this circumstance: the autonomous system might be equipped with sensors that detect specific signatures, they would be unique identifying the characteristics that would be specific to the military objective, in this case, those military vehicles on that road, and then you would have electromagnetic radiation generally not found naturally or amongst civilian objects. The temporal and geographic restriction: if you started looking at that, how is the system going to be used to effectuate the intent of the commander? Is it going to be an operation in place for 30 minutes or 30 days? Again, you're attacking a road to try to secure the airport, so the timing is going to be critical. Is it limited to an operating area over this particular road, or is it covering the entire city? If it's covering just the road itself, certainly you can program it to operate within those certain geographic boundaries. If it's deployed in a limited area, and that's a military objective, and currently this road would be a military objective, then it's analogous to use like any other type of weapon

system like artillery that targets areas of land that qualifies military injection.

One of the questions I think you also have to talk about is intervention. How does the system communicate with the human operator? It's operating with sensors, it's operating with satellites, it's operating to give the commander and the system an operating picture. Does it communicate back with the human operator like every six seconds? Does it talk to the human operator every six minutes? Does it talk to the satellites in the same period of time? How often is it being updated? You'd want to know those types of answers, if it's six minutes, six hours, six days. What is the operator's ability to terminate the engagement? What's built into the system that allows it to terminate the engagement? If the system loses communication with satellites or other sensors, does it automatically shut down and return to base? Those different factors would ensure that it's being operated reliably. Then reliability and predictability: these are determined at various stages of the weapons design, development and deployment process and this helps ensure that the autonomy in the system effectuates the human intention. The key theme that's discussed in the DoD directive is the system would function as anticipated. This means that you're engineering the system to perform reliably, you're training your personnel to operate and understand the system and you're establishing clear machine and human interface. All these factors work together to ensure that the system is going to target military objectives rather than civilian objectives.

Netta GOUSSAC

I'm going to ask Richard to take us through his perspective of the scenario in just one moment, if you allow me just to ask a couple of quick follow-up questions to you, Michael. So, when we're talking about compliance with IHL, of course the target has to be a military objective -and you talked about how that might be something that can be perceived by the system - but we also have to be aware of the risks to civilians and civilian objects that are around the target. When the user of the autonomous weapons system, the attacker or the party to the conflict, doesn't know exactly where and when the force will be applied on the target, how can the user be certain about what's around the target, certainly enough in order to make the proportionality and distinction judgements that are required of them?

Michael MEIER

I think that goes into the military decision-making process and how this is done with the commander. Is it going to be used like any other weapon

system that the commander has? I've advised commanders for over 30 years, and I've never met one that wants less control. Most of them want more and especially over the battlespace which they're responsible for. I think it goes back into the holistic process: the commander has to have trust in the reliability of the system. Before that commander is going to use the system, he will need to be convinced that this system is not going to pick out the civilian jeeps and the civilian systems. You're going to be concerned about the attacks, those military objectives are going to take out the tanks and the armored personnel carriers and that's what you would resolve in the testing process before it gets fielded. So, if the commander is confident in the ability of this system to operate effectively, I think that's how the commander is going to ensure that when he deploys this system it'll be in compliance with the law of armed conflict.

Netta GOUSSAC

So, it seems as though control - the layman's reading of that term - is important not only over the system itself - understanding the system and what it's foreseeable effects will be - but also over the environment of use and being certain about how to minimize the risks to civilians and civilian objects that might be in that environment. That's a good point at which to bring you in, Richard. What's your view on how this scenario would play out?

Richard MOYES

In this scenario and in the next one, which we've seen already, there are various sorts of open questions and uncertainties that are on the table. I feel, looking at this, that there's a capacity to constrain the location and the duration of time over which the system operates. And in many respects, the system is very similar to military systems that are already in existence. It's not necessarily particularly novel as it's structured. It seems to me that that level of definability of the area of functioning and the duration of functioning means that, even broadly as written here, it could be used legally if it were being applied to an area where there was intelligence that there were specific military vehicles that a commander wishes to strike. Putting it directly over those vehicles had a high likelihood of striking those military objectives. That's in a situation where we recognize that in the target profile of the system, as written here, there's some uncertainty whether it will strike military vehicles and civilian vehicles, but if it were being used very specifically and very specifically targeted at identified military vehicles, then that wouldn't seem to me to be particularly problematic. I think that suggests that the system, as a whole, definitely has

the capability of being used within the law. We have both highlighted this idea of geographic and temporal constraints or boundaries. I think those are very important and they need to be affirmed as positive obligations in the use of such systems. Commanders need to manage those parameters in a way that allows a legal application.

A little bit trickier with the system is this idea of the target profile. Michael suggested that that could be rather more sophisticated than as it's written here and we recognize that for certain types of systems, it is possible to have a target profile that is very specifically tailored to the emissions signatures of certain objects - acoustic sensors on certain anti-ship weapons are quite finely tuned to the particular acoustic signature of particular boats. So, in certain contexts, the target profile might be very carefully tailored to individual objects or very specific object classes. However, in the case study as written here, that doesn't seem to be the case. Michael suggested that there would be processes of development and evaluation of these systems. Certainly, I think those are important, but it would seem important that those processes resolve the sort of open questions here. It's not enough for people using the system to know that it will attack what they want to attack, they also need to know what it will attack, that they don't necessarily want to attack, because otherwise they're not working on a realistic understanding of the conditions upon which force will be applied.

So, I think that is the primary consideration here in terms of understanding how that target profile works and knowing what it captures and doesn't capture. All of this is ultimately about understanding the relationship of the system to the context and we don't get any information here about how busy this road is with civilian traffic of the type described, but ultimately understanding the target profile and understanding, controlling the time and space are fundamental to having a manageable ability to evaluate the context and the conditions in the context, and determine what the effects are likely to be.

Something that isn't noted here is what type of force the system uses. When we're talking about autonomy in weapons systems, we're usually talking about the process by which force comes to be applied rather than the specific effects. That's one of the reasons the sort of superfluous injury concerns don't tend to come to the fore because the actual mechanism of harm is often not really discussed. Of course, a commander would also need to know what type of effects the system's going to use. Because if the system detonates with the effects of a 2000-pound bomb, that's different from if it detonates with a shaped charge that disables a vehicle engine.

One of those would be a more efficient and much safer way of operating, but for the sake of the example, understanding the actual force that will be applied would also seem to be important and understanding the number of applications of force that a system can use. Because if a system can apply multiple applications of force to an object or to multiple objects in the environment, if there is a risk of striking at civilian objects, that multiplication ultimately multiplies the risk involved. I think that's a factor that needs to be explicitly understood by commanders in the use of the system.

Finally, I think that it's been useful in the CCW that certain States have highlighted a broad set of points in the development and movement towards the adoption of a system where human control can bear upon that technology and its development. I think those systems generally apply to all weapons. We want weapons systems to function as they're intended to and to be reliable and to be predictable. I think the key issues in relation to autonomy and not necessarily in those surrounding issues is getting a recognition that the duration and the space of the attack needs to be constrained, and that understanding of the target profile and how that target profile has been constructed, need to be in the hands of the users of the system.

Netta GOUSSAC

Michael, did you have any points to follow up based on what Richard just added?

Michael MEIER

No, and I think I agree with most of what Richard said. I think this does go through some of the process. I think other factors that you could do to ensure the protection of civilians - we always have to make sure you take precautions and attack - you could do warnings to make sure that if you're going to engage in attack on this road that the civilians are warned to get off it. There are ways to do that. I think Richard's exactly correct on the munition that the system has, you would certainly want to use the weapons that are appropriate for this. If you're trying to destroy tanks and armored personnel carriers, you would use systems to disable and destroy those versus other types of weapons. But of course, as we said, not every weapon is appropriate for every situation. And that's what comes out through the targeting process and other aspects when you're doing the planning for such an operation. I don't want to imply that distinguishing between civilian and military vehicles is going to be easy by this process or that all

these issues have been resolved, but I think that's certainly what you're going to expect the system to do before you feel it.

Netta GOUSSAC

Thanks. I have quite a few questions, but I'm going to park them for a moment just so that we can maybe examine a second scenario that will tease out even more of these issues and then we can deal with everything at once. Based on the discussion of this scenario, it's clear that constraints on the operation of the system or constraints built into the system itself in its design, but also on its environment of use, and the interaction between the design and the environment, are critical and that weapon systems are not multipurpose in that they need to be fit for the purpose that the commander is using them for. I want to just introduce a second scenario that tweaks a little bit the kinds of issues that we've been talking about.

The second scenario is a different kind of weapons system, and again, I'll read out the scenario for the purposes of translation. So, in the course again of an international armed conflict State A seeks to defend the perimeter of one of its military bases from approaching attacks by State B's forces. The military base is located in a rural area surrounded by a one-kilometer exclusion zone. State A has the option of using a fixed century weapon - a machine gun - effectively. Once activated, the system can select and attack targets within its range without human intervention until an operator deactivates it. The system uses image recognition designed to identify humans holding a weapon as well as infra-red and thermal sensors.

Of course, this scenario is significantly different from the first one as the system used is ostensibly an anti-personnel system in that it's targeting people. Richard, does this raise different considerations from your perspective than the first scenario?

Richard MOYES

Thanks, Netta. I think my first feeling about this scenario is I'm not sure how I would want to be in the military base relying upon their system to defend me, but I think it does raise certain different issues. The time and space issues are dealt with slightly different here. It's a fixed system, it doesn't move around, so that, of course, immediately constrains the spatial area within which it's functioning. It's also suggested that there's an exclusion zone, which means that presumably civilians and civilian objects are being discouraged from entering the area. Of course, we've seen that structure of management adopted in the past in relation to certain types of landmines, where their use has been bounded within marked and fenced

areas that are effectively working to ensure the exclusion of the civilian population.

Now, in practice, those boundaries of exclusion have certainly been questioned by people within civil society as to how effective they are and how reliable they are in the long run as a mechanism for protecting people. But it points to another possible mechanism of managing context of use, which is to exclude civilians from that area. I'm wary generally of pushing the burden onto civilians to not be in the way of weapon systems, but it's interesting that that gets brought up in this context.

Specifically for me, I mentioned before concern around anti-personnel systems in general, I would argue that, and I think I would develop arguments on that in relation specifically to the process of identifying people, identifying human bodies through sensors as targets. That's not so much about straightforwardly humans getting killed as an outcome - because there may be humans in the armored vehicles that we talked about in the previous section - I'm not so concerned to try and argue on the basis of humans being killed as an outcome, but rather that in process terms, there's something disquieting about reducing humans as humans to a set of sensor data that then results in an application of force. As I said before, I wouldn't argue that on the basis of IHL as demanding that as an outcome, but rather in relation to a sort of ethical disquiet and also probably on the basis of a more general societal precaution that we have not generally adopted many systems that function in this way automatically. They, of course, do exist and the capability exists, but it's not widespread to have this level of automatic killing of humans. And I think our societal best interests are probably not served by further encouraging that development or further allowing that development. Of course, individual States may see military advantage in systems that function in this way, I wouldn't deny that. I just feel as a person who doesn't feel particularly advantaged in these situations, that greater movement towards anti-personnel systems in this space would be problematic.

Holding a weapon is not necessarily sufficient for you to be identified as a person who can be legally targeted in these contexts. There are questions about when somebody transitions to the point of being legally targetable. Of course, we have concerns about people who are injured and rendered *hors de combat* as well which play out in this space. I'm trying to avoid really engaging with the sort of more pragmatic questions about how reliable this "holding-a-weapon" notion is? Because these scenarios are sort of theoretical, but holding a weapon doesn't seem to me to be a particularly useful basis for identifying somebody, not least because presumably you

can put the weapon behind your back or there's different types of weapons, it just seems slightly unreliable as a mechanism. If it were a system that was responding to the firing of a gun towards the base, I would orientate to it slightly differently, because it wouldn't involve using the person as a body, as a basis for targeting.

Michael MEIER

This scenario resembles, I think, as Richard pointed out, in sort of existing systems. It tracks similar with what you have with the Korean SGR, the century gun system, that has the century tech system from Israel and obviously, to a degree, it would be sort of a variant of the Phalanx CIWS and C-RAM. Again, as Richard pointed out, the difference in this one is that it's anti-personnel, but I do agree with Richard that there are personnel inside tanks and armored personnel vehicles too. So, those are targeting persons.

I think I do agree with Richard on certainly the image recognition to identify humans holding a weapon. Whether or not that is sufficient for targeting someone, I think it goes back to a lot of the same issues that we had with the other scenario. I would go back to the sort of the intervention by the person: how is this being monitored with other sensors and aspects to understand that these would be combatants? Certainly, the exclusion zone helps with that. Again, if you're in an international armed conflict and it's an enemy combatant, they're targetable holding their weapon or firing their weapon. Certainly, that doesn't stop them from being targeted, but based on this I do think there needs to be more data. I think we would need more data to figure out how this system is going to operate effectively and lawfully and how it is to determine what humans to combat, machines or civilians. Relying on infra-red and thermal aspects may not be sufficient. I think you'd need other safeguards on that one.

Netta GOUSSAC

Sorry, Richard, go ahead.

Richard MOYES

I don't want to throw you off course Netta, but I just wanted to raise another issue which might also lead to some other lines of input, which is just thinking about how these target profiles are built and how image recognition of a person holding a gun might be developed. It was mentioned yesterday, machine learning as a particular sort of technological capability. You could imagine that machine learning might be used to try

and identify images of people holding guns and a system could be trained on that basis, because there's no indication of that in the data here. But I just wanted to raise those specific concerns perhaps about using machine learning to build target profiles. Because with machine learning you can train it on a data set, there are challenges in relation to bias and other factors within that data set. There's also a challenge that you, the human recipient of the system don't actually know what physical characteristics or what image characteristics the system is identifying. You may be able to test it and see that it identifies people holding guns in 90% of the cases, but it's not necessarily possible to identify what the other things are that may result in a false positive identification. In that respect, target profiles built on machine learning are rather different perhaps than target profiles built on a simpler architecture, because the error states are more comprehensible and understandable. I just wanted to flag that because I think if target profile understandability is important, then machine learning raises particular questions in the construction of target profiles.

Netta GOUSSAC

That's exactly the point I wanted to raise. So, here we have a target profile that both of you have identified as being rather too crude, but what if there was a target system that the user was assured could reliably identify a human holding a weapon, but the user didn't know on what basis that identification was made because it was created using machine learning. Would that change your perspective of this scenario? And, in particular, I wanted to ask you about whether supervision would be used in order to counteract that risk.

Michael MEIER

Answering the second part first, I think supervision would probably be very helpful in this situation. I think that's what you have now on the Korean system, and I think the Israeli system as well. Humans do make the decision to actually engage the target.

Netta GOUSSAC

So they're not autonomous…

Michael MEIER

That's correct. They're not fully autonomous, but I think what you get into on supervision goes back to my other point: how often is the human and the machine communicating back and forth? How often does it

communicate with the other ones? So, it may be you don't have the operator engage in the target and it could be human supervised where they can watch and monitor and then turn it off. I think this is the problem you run into with the whole argument on autonomous weapons - technology today will not be the technology you will have 25 years from now. This may be really hard to do today; it may not be hard to do 20 years from now. But I think you do need to have some sort of human supervision monitoring and understanding when this is developed and tested and fielded, how it's going to operate to make sure that you have it. I think just the minimal aspects you have here wouldn't make it hard to field a weapon like this.

Netta GOUSSAC

And if the profile was constructed using machine learning and the user in the field didn't know exactly what factors were being used by the system in order to positively identify a target, would that be a concern in your view?

Michael MEIER

Well, I think again, under the DoD directive, the operator has to understand how the system operates and make decisions. And I think that is what DARPA, the Advanced Defense Research Agency, is doing with explainable AI. So, the operator can understand how the machine makes its decisions, because it's not going to do it in the same way a person does. In fact, even if you ask people how they make certain decisions in certain circumstances, they can't tell you and they'll try to somehow justify it later on. If you ask them about it because you do things without necessarily thinking about it. The same way with the machine itself, it's going to make decisions in different ways. But you want to have a process where you can go back and understand the process that the machine went through to make the decision. And I think explainable AI is a way that can happen.

IV. The use of artificial intelligence in warfare

Artificial intelligence and machine learning: where do we stand and where do we go from here?*

Raja CHATILA
Professor, Institute of Intelligent Systems and Robotics,
Sorbonne University – Campus Pierre & Marie Curie

Thank you very much for the introduction and thank you very much for inviting me to speak today. This is a very impressive audience.

My mission, and I have accepted it, was to give you a crash course on AI in less than 15 minutes. I hope I will achieve this. I will not speak of autonomous weapons; I am speaking about AI in general.

To start with, a question of definition: there are many definitions of AI around and it's better that we are really looking at one definition which I believe is comprehensive enough. What is an intelligent system? I prefer this than Artificial Intelligence. It's a system, a computer-based system, a computational system, which means it accomplishes computations. How? By executing algorithms. It's an organized set of algorithms and, of course, the algorithms are designed by human beings and not by the machines themselves, even if the human beings can use the machines to improve their design. The system is going to use data and these data might be provided to it and might be a very large or a small amount of data or sensed by the system's sensors. And why is an artificial intelligence system or intelligent system built? It's to solve problems, to help us humans solve problems. These problems might be more or less complex, of course, but what we want is to use these systems to solve complex problems and more or less complex situations. I will come back to this issue of complexity which is really key.

The system might have the capacity to improve its performance and this is what we call learning. There are different approaches to learning and I'll mention two main approaches: one is called, today, deep learning, it is actually based on neural nets and I will get back to that as well, or

* The following text is based on the transcript of the recorded session. It has not been revised by the speaker and does not commit him with regard to the views expressed.

reinforcement learning, based on improving the behavior based on previous experience, trying to optimize the best actions.

We sometimes speak about autonomous systems and this is a tricky word as you know. In my perspective, what is an autonomous system? It's a system that is able to accomplish a task, a mission, in a given domain. It has been defined, designed for this domain and for this set of tasks, but it is autonomous because it doesn't require human intervention or human help, to some extent, because it can cope with some changes in its environment. Of course, these changes have to be framed, as it is impossible that the system does anything in any kind of situation, it will always be framed in the given environment. Here are some examples: the issue of intelligence and autonomy, for example, is related to environment and task diversity, complexity, uncertainty. There are a lot of complexities, things are moving in all directions and, of course, we have diverse situations.

This is why you have to define a domain where the system can actually achieve its mission. In the vertical coordinates you have the decision-making capacities, which are related to more powerful algorithms and learning capacities etc. Autonomy is this line, this continuum which starts with automated systems like an automated metro – very simple environment, very simple task, very little decision-making capacities, basically it's automation. As you move into the environmental diversity, complexity etc., you need more and more decisional capacities –and this is why the line is going up. You have the industrial robot, that is, in a more complex situation, that has more degrees of freedom; you have the robot vacuum cleaner, maybe at home, that has to cope with your own apartment; you have a Mars Rover, with a different kind of terrain; and you have the self-driving car on the freeway or highway, because in this situation it's simpler than in the city –we don't know how to actually make autonomous cars in the city today, it has not yet reached technological maturity so that we are able to do it, but we know how to do this on the freeway. If you have a social robot in public spaces, that's much more complex, you have some experimental devices, of course, but this is also an increased level of autonomy. This is the way it goes.

Machine autonomy is actually related to this capacity to determine actions by its own means which means, as I showed, the actual capacity to adapt to a more or less changing and complex environment. In robotics we define two kinds of autonomy: one is called operational autonomy, which is related to the basic actions of the system, like motion, perception, going around obstacles. If it's a flying robot, it's going in a given trajectory through waypoints etc. all by itself. And decisional autonomy, which

relates more to making decisions, devising plans for the future to achieve objectives. This requires more symbolic reasoning, more advanced capacities and today, this is much less possible in robotic systems and AI systems. This is an era of research.

A system has an architecture, it's organized. In this image you have a depiction of an autonomous system. It usually has several layers. Down there you have the basic perception and action capacities that control its sensors, effectors, extract data and control motions. The intermediary level is a control of all those modules, all those algorithms that have to achieve those tasks. At the higher level, you have these decisional capacities which are going to make the plans for the future. This capacity, the planning capacity etc., enables to devise plans for the future based on goals that are usually given by humans. If the human has to control the system, the human can control the system at this level by changing goals, or at this level, or at this level. But here this is real time, therefore, it's like teleoperation. If the human controls the system at this level, it's possible, but you have to be very careful that the internal control system doesn't try to control the system at the same time, because there would be a conflict here -so it's also a problem of design.

Let's speak about learning: basically, learning today is using neural nets. Neural nets have been around for some years. The first mathematical model of neuron dates back to 1943. The neural net operation model dates back to 1956. This is a very simple neural net: it is a set of neurons that interact together and basically you have the input signals here that go to some neurons –for example, this is a set of images, so it's like images coming in- and you have here neurons that are going to make a computation. The neuron makes a very simple computation and outputs value. This value goes to the next neuron or neurons and the importance of the signal that's going out here depends on this W which is a synaptic weight: so the more W is high, the more the output of this neuron is going to influence the behavior of this one and so on.

You have here the so-called output layers that provide the output result and in the middle you have hidden layers simply because they are between the input and the output and if you have many such hidden layers, like a 100, it's a deep learning system. Of course, you can have many, many neurons in each layer: you have neural nets that have maybe 100. That's really large. In each layer you might have hundreds or tens of neurons, this is the size. In our brain there are 100 billion neurons and they are not interacting in a simple manner.

The main issue about neuron nets is that you can use them to approximate almost any function. Here, we are going to use them to learn, to train the net to identify some output that we desire based on the input. This is the principle: you have a kind of output, for example, you want to detect some objects like cars and you have images coming in, so, you have an optimization process designed by humans, so that the synaptic weights are computed to provide the exact expected output. I'm going to explain this using a very simple example, very, very quickly.

Here is a very simple neural net which has 3 layers: input, output, and an intermediate layer. I want to build this net to recognize if there are faces in the images. So, this is not about face recognition, it's just: is this a face or something else? To train it, I'm going to give it a lot, hundreds of thousands of images with faces and others without faces. Here these pictures are going to be processed: you have from these images some small vertical, horizontal, diagonal lines, which are the result of the processing, and you are going to assemble them together until we reach something which is defined as a face, which is here two small lines, these are the eyes, and here one line, which is the mouth. This system will provide an output saying "I recognize the face" when simultaneously you have this and this together. For this system a face is those two eyes and this mouth, so, of course, I can trick it very easily by presenting something which looks like the expected output. Of course, there are a lot of issues in the design of the system that I, as a designer, am responsible for. For example, the training data might include bias, because I trained it only on white males and it will not work properly on other populations. The choice of features that are here, of course, the system is going to detect them by itself, but I programmed the system and trained it so that it's going to select a good choice of features. This semantics, the class semantics, the meaning of these classes –this is eyes, this is mouth and this is face-, I decide that. The architecture must have three layers and all the optimization algorithms that are going to compute those synaptic weights here to provide the expected solution are, of course, designed by myself. And you have issues, research problems and difficulties at all levels.

Here is an example taken from a paper published in April 2019 that expresses some of those limitations in machine learning. The system here recognizes in this image a school bus with a confidence level, a priority of 1. This here recognizes a snow plough with a very high priority as well. This is a motor scooter with high priority and this is a bobsleigh with very high priority as well. Now, what's the problem here? Why did the system do that? Basically, because the learning process is based on this high

amount of data. Actually, you trained it to provide the good solution, but you don't know what it has actually focused on in the images, it's not what you see, it's what it sees. It's really the regularities that it finds in the image. And it might find something completely different focused on something completely different from what you, as a human being, recognized to be such or such object because you actually have this global knowledge that it doesn't have. And, therefore, here, for example, maybe it focused on the snow and probably when you have snow and something transversal and its training set, it was snow plough, and so on. So, in a way, it's very easy to trick a neural net system.

But this is also an arms race, if I may say. Of course, when you discover this, you are going to work and there are a lot of methods and algorithms to cope with that until other issues come up and you have to address them as well.

Now, about autonomy. I would like to really stress the fact that, as you have seen in my previous slides and in this example of learning, the machines really operate at a computational level. It's about computation, mixing data and providing answers to data. The machines don't understand what they are doing, they don't understand the meaning of this data. You understand something about this data to some extent, but the machine doesn't, it just crunches data. So, therefore, its decision, any machine decision, is the result of the algorithm that you have written and that you have put in the machine or that the machine has learned through the whole learning process, which has the same issues. And the machines can only act in the bounded set of decisions that have been provided. Therefore, and I'm insisting on that, machines cannot make ethical decisions. It doesn't know about ethics. Actually, it doesn't know about human beings at all. It knows about computations and those computations represent some concepts for us, but for the machine it's just numbers. They crunch numbers, they provide results, but they don't understand what it's about.

(This is my last slide) Open issues, actually - where are we going in this situation? Research is advancing very fast, so problems of today are going to be solved and new problems are going to emerge. But there are some foundational aspects that will remain. Of course, there is an issue with the training data expressiveness, how much the data expresses, actually what I want the system to learn. Any statistician tells you that if you want to build some knowledge about a population, whatever the population is, you have to have the representative sample of this population. If it's not the case, you will get very wrong results. And this is an issue here about bias. Data-driven machine learning is not contextual, the machine doesn't understand

what it's doing, it doesn't understand the semantics. This is very important if you want to make a decision, it's not just about an image that you are looking at, it's about a whole context, it's about the past, it's about knowledge that you might have as a human, about a situation that the machine doesn't have. Even if it's a sequence of images, for example, a video, will the system operate correctly? Is it dependable, as any other technical system? This is not really today taken into account in the design of those systems because dependability means that you can prove some properties on the system, you can verify it, and with a learning system it's very, very tough to do that.

Robustness. Will the system achieve the expected results despite some changes in the environment? Will it adapt to its environment? It depends on its ability to generalize in its learning system. And this also gets us back to the whole process of learning, training and then testing of the system.

Transparency. The system was built by engineers. Do we know what's inside? Do they even know what's inside? Because sometimes they use software or their training data that comes from somewhere else and it's used as it is. So, there is an issue about transparency, because we need to know what the design concepts are, how the system has been actually built, what the issues are, what values were inside. Is the system predictable? Is it explainable? As I mentioned earlier, we need to know why the system made some decisions, and this is a big issue, it's an open problem. Darko has a research program on that. Can the operator understand why the system has taken certain decisions?

Explicability is at different levels. It depends to whom you are explaining, the competence of the person to whom you are explaining. But at the basis you need to know some things, some data, some processes about how the system actually made its computation and on what, for example, it focused when it considered that this school bus is a snow plough.

Accountability. These are my last words here. I hope I have conveyed the idea that humans are always responsible for the system they build. The system doesn't understand what it's doing and the system includes a lot of issues, that prevent it from understanding what it's doing. So, accountability must remain with the human, of course, the humans who have designed or deployed the system. Therefore, we should ensure that there is a mechanism to identify accountability. Thank you.

The contemporary use of - and possible limits for - artificial intelligence in warfare: a military perspective

Sean MOORE
Assistant Head, Legal – Development, Concepts and Doctrine Centre at the Ministry of Defence, United Kingdom

Thank you for the very kind invitation to take part in this 42nd Round Table, co-sponsored by the International Institute and the ICRC. It is a genuine honour to be amongst such distinguished and knowledgeable speakers. And it a real pleasure to be back in Sanremo. It is far more years than I care to remember since I was a newly qualified naval lawyer getting my first taste if IHL as a student on one of the Institute's courses.

A bit about DCDC

Before I attempt to address the question that I have been given: the contemporary use of – and possible limits for – artificial intelligence in warfare: a military perspective, it might be helpful if I explain a little about what DCDC or, to give it its full title, the UK's Development, Concepts and Doctrine Centre does, and how the law (not just IHL) is at the heart of our thinking.

Bringing together Army, Navy, Air Force personnel, as well as Civil Servants and colleagues from many partner nations, DCDC's outputs and responsibilities include: the Strategic Trends Programme which provides the long term strategic context for policy makers (and if you are interested in our perspective on the possible broader societal impacts of AI, do please read Global Strategic Trends 6); Concepts which outline how our armed forces and defence may operate in the future; and Doctrine, which provides guidance for commanders based on best practice and operational experience. The work of DCDC underpins MOD strategic and joint force development.

As well as contributing to all of DCDC's outputs, my small legal team is responsible for the UK's Joint Manual on the Law of Armed Conflict as well as our guidance and instruction on the treatment of Captured Persons. We also supervise the delivery of IHL training across the Services. And,

particularly relevant to this Round Table's discussions, we discharge the UK's responsibilities under Article 36 of AP1 by conducting legal reviews into all new weapons, means and methods of warfare (and I will say a little more about that later) and a member of my team is one of the UK's Group of Government Experts delegation in Geneva looking at Lethal Autonomous Weapons Systems.

Human Machine Teaming rather than AI or Autonomy

I'd also like to say something about terminology and the deliberate choice of the UK to refer not so much to AI but to Human-Machine Teaming.

Much of what I will talk about today is based on a Joint Concept Note DCDC published last year: Human-Machine Teaming. That choice of words was deliberate. There is, or can be, a tendency to become fixated on the technology around AI, particularly as, if you are not steeped in AI coding, it can seem mysterious and potentially omnipotent. But it is not the technology behind AI that is at the core of the legal and ethical issues, but how we choose to use it.

The developing nature of the technologies in this field has created an array of terms and terminology which are often used interchangeably or differently by various commentators. Drawing distinct boundaries between those terms can often prove difficult, if not impossible. We are seeing some of the practical impact of this complexity in the LAWS GGE discussions under the CCW Convention in Geneva. Those of you who have been following those discussions know that the superficially simple task of agreeing some common definitions has proved a significant challenge in attempts to come up with a framework for the governance of such systems.

And, for clarity, it might be worth me making clear that the UK Government's public ally stated position is that we do not operate and do not plan to develop any lethal autonomous weapons systems.

It is the joining up of humans and machines that is where the contemporary debate about the use of AI is, at least in governments and the military. We are – not yet – in the world of advanced general AI that some like Elon Musk worry about; where human beings become just the biological 'boot-loaders' for a future super-intelligent artificial form of life. And we are not yet in an era where autonomous killer robots go rogue and decide that human beings are an unnecessary inconvenience.

Throughout history, new technologies have been a driver of military adaptation and advantage. Whether moving from sail to steam, horses to tanks, or the introduction and exploitation of the aeroplane or radio, the results have often been transformative.

When it has been transformative, strategy, tactics and technology have often evolved symbiotically; invariably when people figure out how best to exploit the full potential of the emerging combination of technologies.

Robotics, and AI, machine learning, big data, and, in due course, quantum computing and quantum sensing, offer the potential for another inflexion point in delivering military transformation and advantage. However, machines do not yet perform as well as a human brain. So, realising this potential will depend on understanding the relative strengths of humans and machines, and how they best function in combination to outperform an opponent. Developing the right blend of human-machine teams – the effective integration of humans and machines into our war fighting systems – is the key; and we should not forget that we are in a race with our adversaries to unlock this advantage. The clock is ticking, as new technology capabilities accelerate.

Contemporary use of AI in warfare

The array of potential forms taken by remote and autonomous AI enabled systems, and consequently how they interact in human-machine teams, is extremely varied. In size and complexity, they could range from a future AI and robotically-enabled aircraft carrier retrofit, to a single, disposable nano-unmanned aerial vehicle. We typically think of these systems as physical robotic systems in the battlespace, however, applying AI particularly for command and control functions and cyber operations will be increasingly common and important.

Because of the ubiquitous nature of the dual-use technologies of AI and robotics, the impacts on conflict are a matter of when, not if. The effects of these technologies on economics, conflict and society are likely to be increasingly profound and, in the long term, offer new opportunities for strategic overmatch and operational advantage. Harnessing AI will, potentially, give us: increased situational awareness; lighter physical and cognitive loads; sustainment with increased anticipation and efficiency; increased force protection; and, ultimately, superior manoeuvre options in and across all domains. The greatest advantages the confluence of artificial intelligence and robotics development will allow are:

- the ability to scale physical mass and presence on the battlefield independent of the numbers and locations of human combatants;
- extending the reach and persistence of our intelligence, surveillance and reconnaissance (ISR) and weapon systems; and
- information advantage for understanding, decision-making, tempo of activity and assessment.

Command and Control

The use of automation offers opportunities to better exploit information to improve understanding, decision-making and tempo. It will also enable smaller headquarters and more agile command and control. Current UK command systems remain based on significant numbers of staff in static locations with large installed information technology systems. Current configurations are rigid, vulnerable to attack and expensive to reconfigure or redeploy. The move from paper-based to electronic-based workflows has added information awareness and data volume, but at the expense of reduced mobility or structural flexibility. In addition, future intelligence, surveillance, target acquisition and reconnaissance systems will generate much larger volumes of real-time data which will be impossible to process without automated support. Data fusion, automated analysis support and visualisation technologies will be essential to achieving manageable cognitive loads, not just for commanders and staff, but also within individual platforms – warships, tanks, aircraft and, eventually, for individual soldiers, sailors and airmen.

Cyber

The application of AI and automation to cyber systems is the most immediate arena for evolution and advantage. The cyber domain's intrinsically codified nature, the volume of data, and the ability to connect the most powerful hardware and algorithms with few constraints of bandwidth, power access, or limits on speed and repeatability of actions, creates an environment where AI can rapidly evolve and optimise to their assigned tasks.

We must consider that the evolving cyber domain will be a complex ecosystem containing billions of competing AI agents. In the civil sector alone, before any combatant AI systems engage, there will be intelligent

agents competing over: cyber security; finance; media influence; virtual currency mining; advertising; social media influence; pornography; and every other form of web-based interaction. Furthermore, the Internet is dissolving boundaries between the online and physical world. Any deployed cyber system will be exposed to, and become part of, this wider ecosystem; an ecosystem that will also be increasingly indivisible from civil critical national infrastructure.

Remote and Automated platforms

The confluence of AI and robotics development will allow us to scale physical mass and battlefield points of presence increasingly independently of numbers and locations of human combatants. This is similar to the way the Internet has enabled access to information and projection of influence at scale and across the globe by individuals in the virtual domain. Cheap and relatively simple systems are already altering the economics of warfare; an area where the NATO has enjoyed a technological-economic advantage since the 1980s. In March 2017, the US reported that an ally had used a $3 million Patriot missile against a quadcopter that cost $200 from Amazon. Houthi fighters in Yemen, have employed low-cost drones to disable Patriot missile systems in Saudi Arabia. Future options, such as pilot tunnelling, where defensive systems are overwhelmed by employing massed cheap systems, are increasingly viable. Understanding what this means to the way we fight and force development will be significant.

Novel combinations of human-machine teaming will offer a range of new capabilities. They will present opportunities to augment human teams and manned platforms and even create massed effect, such as swarms. Networked mass – large numbers of interconnected sensors and soldiers, vehicles, ships and aircraft – contribute to resilient ISR networks, understanding and enabling manoeuvre. Cheap, smart systems can provide resilience by absorbing casualties on a scale that will not be viable, or desirable, using a solely manned force; they will also be used to overwhelm an opponent's defences.

You will have noticed that, so far, I have talked very much at the conceptual level of what AI and autonomy can do. But the reality is that we are already beyond the conceptual stage.

In 2016, the Royal Navy, working with industry and other partners hosted Unmanned Warrior. In the testing environment of the Scottish coast and waters, for two weeks, over forty organisations demonstrated more

than 50 different unmanned and autonomous air, surface and sub-surface systems.

The Fusion system onboard the F35 aircraft, in operational service today, relies on a high degree of AI in its fusion of remote and onboard sensors. Similarly, the combat information system being developed for the Type 26 ASW frigate will rely heavily on AI in its operation. The Brimstone missile has a long range, seek and destroy mode for when targets are not visible. And close range (and not so close range) naval missile systems, such as Sea Ceptor, have the ability to engage targets automatically.

Possible limits

So, having outlined some of the uses of AI in warfare, I now need to highlight the possible limitations on those uses.

Before I do so, I want to offer a personal opinion about the role of law as a limiting factor. Clearly for those of us who are lawyers, the law, and particularly IHL, is probably the first source of constraint that we instinctively turn to. And, of course, we recognise that law has a normative function – hence the hotly contested debates taking place in the GGE in Geneva, not to mention similar discussions that ended in 2016 over cyber activities.

But the law is, in the short term, not, I believe, likely to be the most significant constraint. Societal, cultural and public opinion will have far more impact on the development of these systems, at least in democratic societies and at least until the acceptance of these technologies in everyday life becomes widespread. Most of us will be familiar with Google's ending of its relationship with the US DOD after its employees objected to being involved in the business of war. And it is public opinion more than law, that lead the UK to make public its no development of lethal autonomous weapons policy.

Added to public opinion, must be a healthy degree of scepticism over the technology itself. Any of us who have served in uniform will know that even the most expensively procured piece of equipment rarely stands up to the rigours of military life for long. And often fails at exactly the wrong moment. Never underestimate the ability of the ordinary solider, sailor, airman or marine to break things {just ask the British Army how many Watchkeeper drones they have lost]. And that is without the interference, spoofing, jamming or deception of our adversaries. And, whilst I would be

the first to acknowledge the impressive speed of development of AI, it still remains fragile. So, it would be a foolhardy government that bet the farm on AI.

So, with that caveat, I will mention a few of the legal issues that we as military lawyers are wrestling with as we look at AI and warfare. And as operational lawyers, I should say that IHL is only one field of law that governs the conduct of military activity. Whilst IHL remains central, the modern military lawyer needs to be able to advise his or her commander on a range of legal issues, from IHL to human rights law, domestic and international criminal law to data protection.

Article 36

Central to the UK's position in Geneva has been that, whilst in theory turning the battlefield over to killer robots might be an attractive notion to science fiction writers, doing so will not relieve States of their obligations under IHL. And, like any technology deployed in warfare, that must be capable of being used in compliance with our international obligations: whether it is a dumb bomb or a smart robot.

As I said in my introduction, it is my team at DCDC that conducts Article 36 reviews on behalf of the UK. And in doing so, we must ask ourselves can this weapon be used in compliance with international humanitarian law. If the answer is no, we do not pass the system.

I won't repeat what both Professor Dinstein and Professor Nasu said yesterday about discrimination and superfluous injury or unnecessary suffering. But these key principles are at the core of the Article 36 process. So, when looking at a future autonomous weapon system, we would have to direct ourselves to the very heart of what drives the fear against such weapons. Put simply, an autonomous weapons system that was incapable of being used in a proportionate and distinctive way, or a weapon system that risked causing unnecessary suffering or superfluous injury would not get near the battlefield.

It is true, however, that the sophistication of new weapons, particularly highly automated ones, with sophisticated algorithms, incorporating machine learning, does pose new problems for those of us conducting such reviews. But those problems are not insurmountable.

The F35 aircraft is by no means a lethal autonomous weapon system. But it does possess a high degree of autonomy, and the sophistication of its combats system that fuses together data from a range of onboard and off

board sensors is beyond anything previously seen. Therefore, whilst it is the pilot who makes the decision to fire a particular weapon, in reviewing the aircraft, we had to make sure that he was not simply a rubber stamp to a decision that had in fact been presented to him by a machine. We did this in the way we would any weapon, by looking at the evidence of trials and analysis by systems experts. Conceptually the approach was no different even though in practice the volume of data was far in excess of anything we had previously done.

Article 82

Notwithstanding what I have said about the importance of Article 36, I do think that the nature of machine learning based systems, with increasing levels of autonomy means that the traditional boundary between the obligations under Article 36, i.e. before a weapon is deployed, and Article 82, the duty to provide legal advice during operations, will become blurred. As systems become capable of operating independently for longer periods and capable of self-programming and adapting, the need to monitor these systems and provide legal advice to the commander will remain. Whether through supervisions by humans or by programming in 'check if in doubt' protocols into the systems, will mean that the need for legal advice will become ever more valuable.

IHL Programming

Whilst I fully agree with what Professor Dinstein said about being a long way off from machines that make proportionality decisions, we are already able to programme certain ROE parameters into some systems, for example, that do not engage certain IFF responses.

A point that Professor Dinstein alluded to yesterday is worth drawing out. Autonomous weapons systems could, if programmed properly be more compliant with IHL. A low cost, disposable weapon system does not need to be taught the virtues of courageous restraint. It could be programmed not to open fire if there is any doubt as to the legitimacy of the target. Any reasonable commander would be quite content to see a cheap drone destroyed by the enemy rather than cause incessant civilian casualties in a way that they would not for any man or woman under their command.

I said a moment ago that operational lawyers have to think broader than IHL to do their job, so let me conclude by pointing out some of the other aspects of law that can and will limit the deployment of AI systems.

Status: there is a significant debate underway within the naval law community over the status of autonomous vessels. Are they warships with the right to engage in attack, or are they auxiliaries, or are they merely weapons systems? The reality is that they will be a mixture of all three. But uncertainties over status do limit – in the short term – the development of such systems.

As we heard yesterday from Professor Gaggioli, human rights law is playing an increasing role in some armed conflicts. And, in the quasi-NIAC conflicts we may be facing, where the threshold of IHL applicability is sometimes blurred, and our adversaries are willing to use a range of activities of varying legitimacy to achieve their ends, then human rights concepts such as privacy and the right to life are, in some circumstances, relevant considerations.

More prosaically, we should not forget that these systems are designed by engineers, and if there is one group of people who love rules and process more than lawyers it is engineers, so whilst not necessarily hard law, we must not underestimate the impact of regulatory standards and basic health and safety processes in ensuring that unpredictable systems are not released into the wild.

And as a concluding remark, I could not agree more with Professor Venturini when she said yesterday that training is the key. Even more than law, educating our people in the merits and the limitations of these systems is essential. That does not mean that we need to create a generation of computer programmers. For thousands of years we have lethal autonomous weapons systems, they just happen to be made out of flesh and blood.

Thank you for your time.

Artificial Intelligence in military decision-making: which limits does IHL impose regarding targeting and deprivation of liberty?*

Heather HARRISON DINNIS
Senior Lecturer, Swedish Defence University

I have been asked to reflect on what limits international humanitarian law (IHL) places on two specific areas in which technologically advanced militaries have been investigating the use of artificial intelligence to assist military decision-making, that is, of targeting and detention.

I use the word 'assist' advisedly. Algorithms cannot (and in my opinion should not) replace human decision-making in such matters. However, they can provide a valuable tool by doing what algorithms do best – sifting through vast amounts of data and establishing patterns and statistical probabilities. The previous speakers have talked about the current state of technology [and the uses to which it may be put in an armed conflict situation] so I will take that as read and focus my comments accordingly.

I'll begin with two general points of note regarding IHL considerations that relate to both targeting and detention issues.

First, and for the avoidance of doubt, algorithms that are designed to assist military decision-making (including those which employ machine learning) are subject to the same rules as every other military technology. Further, depending on the purpose for which they are deployed, and the accompanying IHL obligations of the State employing them, they may also be subject to AP1, Article 36 review (or equivalent) either as part of a weapons system or as a method of warfare. The nature of algorithms being what it is, this review process may need to take place on multiple occasions with far more regularity than conventional weapons systems, and the particular challenge of machine-learning algorithms will need to be accommodated in some fashion.

The other general point relates to a now relatively well-known problem with the use of algorithmic decision-making tools as they are used in the domestic context, that of bias**.** It's one that that State armed forces and

* The following text is based on the transcript of the recorded session. It has not been revised by the speaker and does not commit her with regard to the views expressed.

military commanders specifically will need to be particularly aware of in the high stakes context of armed conflicts (or other crisis situations).

Customary IHL contains a general prohibition on adverse distinction or discrimination on the basis of race, nationality, gender, etc.[1] But algorithms are merely computer codes – they reflect the biases of the programmers who write them, the biases inherent in the data sets that they are trained on and may also be used in an overtly biased manner – for example, to search for a particular minority or ethnic grouping. Studies have shown that there is a significant racial and gender bias in the algorithms currently being marketed and used in western democracies by law enforcement but also in the criminal justice system, for example, when deciding whether to release a defendant on bail pending trial, or grant parole following detention. How much more so would these biases occur in algorithms either transplanted into an entirely different cultural situation and context, or where insufficient data sets exist (or are available) to the military wishing to use them – particularly at the start of a conflict. Ashley Deeks in her paper, Predicting Enemies, has pointed out that had the US begun collecting information on detainees in Iraq and Afghanistan at the start of the conflict they would have established a significant data set, capable of providing useful information about detainees returning to the fight.[2] However, in the absence of protracted information collection of a high quality, the chances of bias in any data set used by a military algorithm are significantly higher than in a domestic context where the variables are largely known.

Care must be taken, therefore, to ensure that these biases are minimised to the greatest extent possible at every level. Commanders must be aware of the problem and, therefore, be able to mitigate their own automation bias and the tendency to believe a machine over their own judgments! Inherent (or overt) biases are not translated into unlawful discriminatory behaviour.

I'll turn now to the IHL limitations of the two particular areas we are covering today. I want to begin first with issues raised by detention.

Detention

In an international armed conflict, deprivation of liberty can occur in a number of circumstances. For example, a party to the conflict may detain

[1] Rule 88: Adverse distinction in the application of international humanitarian law based on race, colour, sex, language, religion or belief, political or other opinion, national or social origin, wealth, birth or other status or on any other similar criteria is prohibited (ICRC CIHL Study).

[2] Deeks, Predicting Enemies, 17.

prisoners of war, civilian internees and security detainees. Each of these categories has extensive protections set out in the Third and Fourth Geneva Conventions respectively. In non-international armed conflicts, the party may be detaining those who have directly participated in the conflict, those who are detained on criminal charges, or those who are detained for security reasons related to the conflict. The treaty law protections for these detainees are more general in nature (Common Article 3 and CIL) but nevertheless raise important issues in respect of AI assisted decision-making.

In terms of AI assisted decision-making, POWS raise the least issues. Their status-based detention and the ability of a party to detain them, without review until the end of the conflict mean that the types of predictive algorithms used in detention situations are largely unnecessary. However, where a person's detention is based on their threat to the detaining party or to the civilian population, issues may arise.

A party to an armed conflict may intern civilians on the basis that they represent a serious threat to the security of a party to the armed conflict (the detaining authority). For example, protected alien civilians in the territory of a party to an armed conflict can be detained 'only if the security of the detaining power makes it *absolutely necessary'*. An occupying power may intern protected persons if necessary, for '*imperative reasons of security'*. In both these cases the use of predictive algorithms, to determine whether or not someone constitutes a security threat has parallels with the use of AI decision support in domestic criminal justice systems to determine whether to release someone on bail pending trial or on parole or early release from a prison sentence. Because the decision to detain must be based on the security threat posed by the individual detained, there is an immediate problem when you are using a data-set based on the previous actions of others. The ICRC's commentary to the provisions notes that 'there can be no question of collective measures: each case must be decided separately.'[3] [To a certain extent this can be addressed by noting that the predictive algorithms work on statistics, therefore, they mimic the unconscious processes already used by the judge's experience]

Certainly, merely transplanting an algorithm designed for a domestic setting to an armed conflict setting will not be a viable solution. The definition of what will constitute a security threat in each of those contexts is entirely different. But for the military to build their own, in a complex

[3] ICRC Commentary to Art. 78 GCIV on security measures (see also Art..s 41&42).

and often rapidly changing environment, would take a huge amount of resources to establish the detailed information that would be required to build and train an algorithm that could work to the appropriate level of accuracy. The military is interested in accurate predictions that provide security to their forces. The use of AI would be most useful in protracted conflicts where this sort of information can be accrued over time.[4] Militaries will also need to be acutely aware of the inflated claims of some AI providers. Some of the analysis capabilities claimed by AI companies out there are simply based on flawed science. For example, one particular company (Affectiva) claims to measure 'complex and nuanced emotional and cognitive states from face and voice'. There is a reason that polygraph (lie detectors) are not admitted as evidence in most legal systems. There is no way to reliably correlate physiology and external movement with a person's internal mental state.[5] As other studies have noted – how people communicate feelings varies substantially across cultures, situations and even within a single situation.[6]

Review of detention decisions must be carried out periodically by an impartial and objective authority that is authorised to determine the lawfulness and appropriateness of continued detention. Whether such a review could be carried out by means of an algorithm, where the circumstances on the ground may change rapidly over time and there is not enough stable data for the system to work accurately, would militate against using algorithms in these circumstances.

There is also a requirement that the detainees be promptly informed of the reasons for their detention in a language that they understand. The risk of AI assisted decision-making is that an algorithm deems someone a security threat for reasons that are not known even to the operator. The 'computer says no' scenario is of particular concern with machine-learning algorithms which may make associations based on past behaviours, and behaviours of others, which differ from those that a human steeped in cultural cues would judge irrelevant. Further, in the domestic context, companies providing similar services to criminal justice systems have refused to release details of their proprietary algorithms to detainees

[4] E.g. Israel IDF / Hamas. Deeks – info types – tribal relations, neighbourhoods, placed fighters live, loyalties and associations, suspicious travel routes, and enemy tactics and techniques.

[5] ACLU The Dawn of Robot Surveillance

[6] Barret et al 'Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements', cited in ACLU, The Dawn of Robot Surveillance

wishing to challenge the reasons for their detention. It seems likely that similar problems would face detainees in military detention with the added difficulties of the national security concerns of the detaining State and military classification systems making access to the reasoning behind the decision even more complex.

Targeting

The second area where AI assisted decision making raises particular issues for IHL is where it is used to assist the targeting process. I know that our next speaker [Andrea Farrés Jiménez] will be addressing this in more detail, so I will just highlight a couple of examples of the IHL issues where AI assisted decision-making may be used as part of the targeting process.

Again, many of the concerns raised in respect of this issue relate to the interpretation of, for example, culturally based differences between populations. Similar concerns have been raised in relation to so-called pattern of life targeting. How much more difficult it is to ensure that the computer scientists responsible for programming the algorithm and the algorithms themselves reflect the appropriate indicators necessary for compliance with IHL targeting obligations for a specific conflict (particularly at the beginning of that conflict). Concerns might be raised again by cross-cultural differences in respect of indicators of direct participation in hostilities, for example, weapon carrying in Afghanistan (or Yemen) would not necessarily provide a valuable indicator of DPH in a society where it is commonplace for men to carry weapons.

The IHL obligation for commanders and those who decide upon an attack to take all feasible precautions in attack will also be affected by the use of AI assisted decision-making. In situations where the algorithm is essentially a black box, there will be a particular need for commanders to verify that the target is a military objective rather than merely relying on the algorithm to assess. The move towards ‘explainable AI’, that, is, AI that can explain its reasoning, should go some way to address this concern. However, States employing the technology must be acutely aware of the limitations of the particular technology at their disposal and fight against their own automation bias in order to meet their IHL obligations.

That awareness must also extend to the recognition that algorithms analyse and ‘reason’ very differently from human beings and, as one

researcher recently put it, in relation to some forms of analysis, algorithms cannot tell the difference between a polar bear and a can opener.[7] A similar awareness of the difference between humans and algorithmic decision-making will be required when a choice of targets is available., For example, determination of a military advantage by a machine-learning algorithm may introduce elements of machine strategy, which we know from examples (alpha Go trials – distinctly non-human strategy) or in the very contextually-sensitive proportionality calculations required by the targeting process.

All of which is not to say that AI assisted decision-making cannot be used in compliance with international humanitarian law. It is clear that the ability of AI to collect and analyse vast amounts of data may well assist humans in making better decisions in order to conduct hostilities in compliance with IHL. However, States choosing to employ such techniques must be fully aware of the possible pitfalls, the limitations and the risk of misuse of the technology and guard against them at every stage of development, acquisition and deployment. They must be prepared to allow the human to remain central to the process of warfighting in order to maintain that delicate balance on which IHL rests.

[7] Wire report – relating to recognition on the basis of outline.

Presentation of the winning submission to the 2019 Sanremo New Voices in International Humanitarian Law essay competition: "The SKYNET programme and the principle of distinction: why we should not let artificial intelligence lead the way"

Andrea FARRÉS JIMÉNEZ
Humanitarian Policy Intern, Norwegian Refugee Council

Three years ago, big data analytics for pattern recognition in intelligence data came to light through the Snowden leaked documentation. It was the SKYNET Programme, a machine-learning algorithm developed by the United States.

The aim of this programme was to analyse the cellular network metadata of millions of people in Pakistan, to identify couriers carrying messages between Al-Qaeda members, rating their likelihood of being terrorists. Whether the SKYNET Programme was actually used, and the exact characteristics of it, is unknown by the general public.

However, when analysing the leaked documentation explaining how the algorithm functioned, we came across an interesting case study to take as an example to analyse some of the challenges artificial intelligence (AI) can face to ensure compliance with the principle of distinction.

Therefore, the aim of my presentation is to take the case study of the SKYNET programme as the starting point, through which I would like to raise several general challenges of using AI as a decision aid system in relation to the compliance of the principle of distinction in the targeting process.

Starting with how the training of the algorithm was made, I am going to raise general concerns on some difficulties AI faces to ensure the protection of civilians. Also, I am going to tackle how the variants inserted can jeopardize cultural sensitivity and perpetuate biases.

I am also going to raise the practical problem of the scarcity of available and reliable data, and the problems which can result from that. Finally, I am going to discuss the psychological impact operators experience in human-machine partnerships, and how this can potentially impact in terms of IHL compliance.

Before starting the discussion on the SKYNET programme, I would like to talk about some background considerations relating to AI. AI is a branch of computer science that deals with the simulation of intelligent behaviour in computers. Especially during the last years, the development of AI has increased a lot. Currently, it impacts almost every aspect of our lives, for instance when we use our smartphones or when we get Amazon recommending to us what to buy or Netflix suggesting what to watch.

As an essential part of AI, algorithms are mathematical instructions which tell computers what to do. One widely used type of algorithms, also used for the SKYNET programme, are the machine-learning algorithms. Simplifying a lot, machine-learning algorithms are systems which process massive amounts of data to identify autonomous rules or patterns. Through a learning process, they can end up drawing general conclusions from single pieces of information.

This learning process consists of 4 steps: training, testing, application and validation. From now until the end of the presentation, I will explain the concerns, both legal and practical, that arise throughout the machine-learning process the SKYNET programmers could have faced.

The first step to develop the SKYNET Programme was to train the algorithm. The training phase consists of feeding the computer with huge amounts of data labelled as a “ground truth”. This “ground truth” is based on pattern recognition, using mathematical methods to find relationships in the sensory data.

At the outset, it is worrisome that as a “ground truth” of the training algorithm the obligation of refraining from targeting civilians receives no consideration, or at least this is what can be inferred from the leaked documentation. Ensuring that the civilian persons and objects are protected is crucial, as the principle of distinction requires parties to an armed conflict to distinguish between civilian persons and civilian objects, and combatants and military objectives, as only the latter can be targeted.

However, even if programmers wanted to make sure that the algorithm designed would comply with the principle of distinction, it seems that it is extremely hard to reach this result. This is because Article 50 of the Additional Protocol I to the Geneva Conventions describes the category of civilian population in a negative sense. This formulation entails an AI-related challenge which I consider difficult to overcome. Since a definition of the concept of the category of “civilians” is absent, it seems that it becomes very complex to translate this notion into a computer code.

I would argue that feeding the machine with a definition of what is a combatant, and that all what does not fit into that definition should not be

targeted, may not be enough basis to ensure civilian protection. I believe that this is so, because this reasoning forgets a third category of people: civilians directly participating in the hostilities (DPH).

Inserting in an abstract way what is a civilian directly participating in the hostilities can entail various challenges. Those difficulties are due to the fact that, to assess if a civilian is DPH, the threshold of harm, direct causation and the belligerent nexus need to be proved.

And to analyze these 3 characteristics, I think it is necessary to assess contextual information, "the big picture", a task only humans can undertake. For instance, commanders should assess "the tactical and strategic implications of a potential harm; the status of other potentially threatened individuals; the direct causal implications of someone's actions; or the sociocultural and psychological situation in which that individual's intentions and actions qualify as military actions".[1]

Therefore, contextual information, "the big picture", essential to properly identify civilians DPH, becomes a task only humans can undertake, and which the algorithm most likely would leave out of its assessment.

Finally, IHL states that the presence of military or civilians DPH among the civilian population does not deprive the population of the protection from an attack. I would argue that this provision favours the need for military commanders to issue context-based decisions, which reliance on AI computer-aiding would likely omit as well.

Moving on to the "ground truths" which were inserted in the training of the SKYNET algorithm, more issues of concern arise. For instance, the more than 80 properties entered as relevant data in the SKYNET Programme to help rating people as couriers assumed that their behaviour differs from the rest of the population and those variants included factors as turning off the phone or swapping SIM cards, understood as attempts to evade mass surveillance.

The assumption that couriers portray a distinct behaviour in relation to the use of their phones can be by itself problematic. This shows that the process of data interpretation is not neutral, as the biases of the programmers can be reflected in how and which information is introduced to the machines.

---

[1] Peter Asaro, "On Banning Autonomous Lethal Systems: Human Rights, Automation and the Dehumanizing of Lethal Decision-making," Special Issue on New Technologies and Warfare, *International Review of the Red Cross* 94, no. 886 (Summer 2012); 789.

To gain a more accurate insight on which variants are actually relevant or not, situational awareness and cultural sensitivity is needed by those operators. A good example of this need is in conflicts like Yemen or Afghanistan, where civilians carry weapons for self-protection. In these cases, programmers in the US should be aware that carrying guns or not is not a valid criterion to distinguish civilians from combatants, contrary to what someone working in a robotics laboratory in the US may assume at first.

For example, there have been cases of signature drone strikes where civilians were targeted after feeling forced to provide shelter and food to militants in their homes. In cases like this, it has been debated that partially due to such loss of situational awareness, the duress the civilians experienced was not considered.

In conclusion, I would argue that cultural sensibility is an important component for an algorithm to be accurate, so AI programmers should receive specific training on that.

Another challenge I would like to point out is the scarcity of available and reliable data. Machine-learning algorithms need massive amounts of data to infer accurate patterns. This means that feeding the machine with information of dozens of known couriers is not enough.

This is so because the least information available, the more the machine can produce false positives and false negatives. False positives occur when someone is mistakenly identified as a terrorist or combatant, and false negatives refer to the contrary. Besides, the reliability on the information these databases provide depends on their accuracy, which is challenged as sometimes those are not updated or contain mistakes.

As the last point concerning the training of data, I would also like to highlight that in the leaked information the data is not trained to identify *hors de combat*. This could lead to IHL violations if the eventual targeting decision does not consider this option, because assuming that somebody is a combatant without the possibility of contemplating their surrender is an IHL violation.

Once the training phase is done, the testing period starts. This second step consists of inserting another data set to check whether the machine can properly generalize. If it can do so, then it is put into operation, for the third step, which is the application phase. During this stage, the machine analyses a wide range of information to find the patterns that match with the training set. After this procedure, the validation phase starts, in which programmers assess the machine's performance.

The information available on how the testing, validation and application phases were made in the SKYNET programme, spotlight some general concerns.

First, the question of the scarce data available. As mentioned, for machine-learning algorithms to work, massive amounts of data need to be inserted. If all the relevant information regarding known couriers' database is inserted in the algorithm for the training phase, it means that there is no other separate and different dataset available to corroborate how accurately the algorithm works. In this sense, the training phase would get totally invalidated.

In relation to the validation phase, some worrisome concerns arise, as we can observe the psychological impacts operators are likely to experience when working through a human-machine partnership. As the leaked information reveals, what appears to be shown as a proof of accuracy is that the person who got the highest likelihood of being a courier is, in fact, a well-renowned Al Jazeera journalist, meaning, a false positive.

A way to explain these low validation standards is the demonstrated psychological impact technology has on human operators. Indeed, partnerships human-human and human-machine do not work in the same manner.

Human performance is affected by automated systems, having an impact on the "loss of situational awareness, complacency, skill degradation, and decision biases."[2] Operators, and humans in general, have a tendency to over rely on the outcome the computer produces (who has not reluctantly followed a suggested google maps route, even if in fact he or she thought that there were alternative faster ways to reach the desired destination?). Indeed, humans tend to delegate too much to automation. Eventually, this means that the decision-makers are less attentive and healthy scepticism over what the decision aid suggests is erased.

This can have negative implications on some IHL obligations. For instance, it can lower the presumption of the civilian status in case of doubt. Humans tend to over rely on the outcomes of machines, downplaying hesitations and suppressing doubt over what computers suggests. The practical implication of this reaction is jeopardizing the application of the mentioned IHL presumption. Nevertheless, a way to mitigate those impacts human-machine partnerships experience, could be with specific training.

---

[2] Mary L. Cummings, "Automation Bias in Intelligent Time Critical Decision Support Systems," *American Institute of Aeronautics and Astronautics* (2004): 2.

To conclude, I would like to highlight some general challenges taken from the SKYNET example regarding the compliance of the principle of distinction and the reliance of AI as a decision support aid.

Modern asymmetric and urban warfare entail high levels of controversy regarding who can be lawfully targeted. With the fog of war getting thicker and thicker, commanders and politicians are naturally inclined to search for tools to get guidance on whom they can lawfully target. However, AI should not be a substitute of combat-experienced human judgement. The principle of distinction is highly complex, contextual, and requires a kind of analysis only human minds are suited to fully undertake.

The SKYNET Programme spotlights some relevant AI-related challenges, such as the risk of a lack of ensured protection of civilians, or a problematic selection of "ground truths" to feed the algorithm with. The scarcity of reliable data, and the psychological impacts human-machine partnerships have on the human operator, do not seem to help either in lifting the fog of war.

Therefore, when AI is used as a decision aid, what the algorithm leaves out, or what data is considered relevant, is information which must be kept in mind by the military commander, before issuing any targeting decision, so all the necessary considerations to comply with the principle of distinction are taken into account. Indeed, in my view, ensuring the respect of the principle of distinction necessarily requires a human's assessment, which AI is not "intelligent" enough to replace.

Thank you.

V. IHL and challenges related to outer space warfare

Military Use of Outer Space: A U.S. Perspective*

Simone V. DAVIS
Lt. Colonel, Chief, Air and Space Division,
Headquarters US Air Force (USAF)

United States Space Missions

The United States (U.S.) Department of Defense (DoD) - the Executive Branch agency responsible for the U.S. Armed Services - has five core space missions which guide U.S. space operations. The first of these missions is Situational Space Awareness, which concerns the collection of information to help commanders gather as much information as possible from all available sources. This does not occur in a vacuum, however, as the DoD relies heavily on allied and commercial partners to provide situational awareness on U.S. and partner nation space objects.

The second DoD core space mission is Battle Management Command and Control. These are the tools used to give commanders real-time information so they can make on-the-spot tactical decisions. The mechanisms used to execute this mission include satellite communication and GPS technology, for example, that can feed data to aircraft and maritime assets, providing commanders invaluable information to assist in their tactical decision-making.

The third DoD core space mission is Support. How do we actually get space objects into orbit? There are several support systems involved in the process, including space lift, which is the delivery of satellites, payloads, and other materials into space to effectuate space missions, and space operations, which is what occurs once space objects are in orbit. We have to track our objects to ensure they remain healthy and conduct maintenance if they break. Support also includes the reconstitution of space objects. If a satellite being utilized for one of our mission sets becomes disabled, we must have alternative means to be able to accomplish the mission. We have plans and operation orders that guide how the United States will operate in the event our space capabilities are degraded or destroyed.

* The thoughts and opinions expressed herein are solely those of the author and do not necessarily reflect the official position of the United States Air Force or the Department of Defense.

The fourth core mission is the stalwart mission—Space Support to Operations. This mission includes the support we provide to land operations as well as naval and cyber operations via space. This is accomplished primarily through space-borne intelligence, surveillance, and reconnaissance operations. It also includes missile warnings; satellite communication; positioning, navigation, and timing (PNT); and environmental monitoring.

The final DoD core space mission is Space Control, which can be divided into two subsets. The first is Offensive Space Control and is defined as the actions the U.S. may take in order to ensure our adversaries do not negatively affect U.S. space assets. There are five means for accomplishing Offensive Space Control, affectionately referred to as the "Five Ds." They include deception, disruption, degradation, denial, or destruction of hostile space capabilities. On the opposite front is Defensive Space Control, which includes the passive and active acts taken to protect space objects from maneuvering a satellite into a different position or installing additional firewalls to protect the integrity of a satellite.

As mentioned at the outset, these five missions constitute the five core charges of U.S. space operations. Any space-based U.S. activity can be characterized into one or more of these categories.

Sources of U.S. Space Law

Turning now to the sources of U.S. space law and policy. As with all space law, U.S. space law is grounded in several international treaties to include the Outer Space Treaty, the Liability Convention, the Rescue Agreement, and the Registration Convention. Domestically, the overarching U.S. space doctrine is Joint Publication (JP) 3-14, *Space Operations.* This publication outlines how joint space operations will be conducted, including the command and control structure as well as how the DoD interacts with other government agencies. In addition to JP 3-14, DoD Directive 3100.10, *Space Policy*, summarizes the roles and responsibilities of the respective combatant commands and military departments with respect to how each contributes to space operations.

Additional sources of U.S. space policy include the *National Space Strategy* and Space Policy Directives, which are all Executive-level White House space policy declarations. The current White House administration has placed renewed focus on U.S. space capabilities. Most apparent was the issuance of Space Policy Directive-4 in October 2018, which directed the

establishment of a sixth branch to the Armed Services, the U.S. Space Force.

Finally, we have several federal regulations affecting U.S. space law and policy, including those related to PNT and the U.S.'s commitment to provide free GPS to all citizens, as well as, regulations concerning commercial remote sensing and the declaration that we must, to the greatest extent possible, team with commercial partners to use remote sensing.

Space Law and Permissible and Impermissible Space Activities

In accordance with the Outer Space Treaty, the exploration and use of outer space is the province of all mankind and there is a recognition that outer space will be used for peaceful purposes, which has been interpreted as non-aggressive actions. That does not mean there can be no military uses of space, only that such uses shall be non-aggressive. Non-aggressive military uses include intelligence collection, military ballistic early warning systems, satellite communications, and GPS-based navigation.

Additionally, international law prohibits the militarization of the Moon and other celestial bodies to include the building of bases or forts on the Moon and the conducting of weapons testing or execution of military movements and maneuvers. Finally, with respect to the use of outer space, no weapons of mass destruction can be placed or stationed in outer space, however, that does not prohibit the transiting of weapons through space, for example, in the case of an intercontinental ballistic missile, that maneuvers through space but does not stay in the Earth's orbit.

As Article 3 of the Outer Space Treaty notes, international law applies to activities conducted in outer space. But what might that look like? For example, you have Article 2(4) of UN Charter which states that all Charter members "shall refrain in their international relations from the threat or use of force…." Additionally, Article 51 of the Charter acknowledges States' inherent right of self-defense. What would be considered "force" in outer space enough to trigger Article 51 protections?

And what if a conflict were to occur in space? The Law of War would undoubtedly apply but how? At this juncture, we do not necessarily have the answers to these questions, but they are important thoughts to consider as States continue to improve their space capabilities.

Limits imposed by outer space law on military operations in outer space

Elina MOROZOVA
Head of International Legal Service,
Intersputnik International Organization of Space Communications

Ladies and Gentlemen,

First of all, allow me to warmly thank the International Institute of Humanitarian Law for kindly inviting me to take part in this Round Table. It is an honour for me to address you all, the experts who are truly the best in their field.

Within the framework of this session I would like to discuss some topical aspects related to the limits imposed by international space law on military operations in outer space.

International Space Law

Let me start with a brief reminder that space law is generally associated with five United Nations treaties. The Outer Space Treaty,[1] being the first and the most comprehensive, provides a general framework for the regulation of space activities. It is on this foundation that relevant provisions are further developed by the other four UN space treaties. They are the Rescue Agreement,[2] the Liability Convention,[3] the Registration Convention,[4] and the Moon Agreement.[5]

---

[1] *Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, Including the Moon and Other Celestial Bodies*, 27th January 1967, 610 UNTS 205 (entered into force on 10th October 1967). As of September 2019, 110 ratifications, 23 signatures, and 1 declaration of the acceptance of the responsibility for compliance with the Treaty.

[2] *Agreement on the Rescue of Astronauts, the Return of Astronauts and the Return of Objects Launched into Outer Space*, 22nd April 1968, 672 UNTS 119 (entered into force 3rd December 1968). As of September 2019, 98 ratifications, 23 signatures, and 3 declarations of the acceptance of the rights and obligations under the Agreement.

[3] *Convention on International Liability for Damage Caused by Space Objects*, 29th March 1972, 961 UNTS 187 (entered into force 1st September 1972). As of September 2019, 97 ratifications, 19 signatures, and 4 declarations of the acceptance of the rights and obligations under the Convention.

[4] *Convention on Registration of Objects Launched into Outer Space*, 14th January 1975, 1023 UNTS 15 (entered into force 15th September 1976). As of September 2019, 69 ratifications, 3 signatures, and 4 declarations of the acceptance of the rights and obligations under the Convention.

[5] *Agreement Governing the Activities of States on the Moon and Other Celestial Bodies*, 5th December 1979, 1363 UNTS 3 (entered into force 11th July 1984). As of September 2019, 18 ratifications and 4 signatures.

The UN space treaties are supplemented with a number of non-binding instruments – Resolutions adopted by the UN General Assembly and documents produced by the UN Committee on the Peaceful Uses of Outer Space.

Both 'hard' space law and 'soft' space law constitute *lex specialis* which, along with international law in general, governs all space activities irrespective of their nature, while military space activities have always been in the focus of the interest of each State.

Peaceful Uses of Outer Space

At the beginning of the space era, when the first artificial satellite was launched,[6] States realized that outer space had just acquired a new practical value – that was the ultimate height ever reached by humans which could offer significant strategic benefits to the firstcomers. More so, at that time both the Soviet Union and the US successfully demonstrated their nuclear capabilities, and that influenced the formation of space law.

That is why, the UN General Assembly immediately adopted a Resolution[7] urging States to ensure that the sending of objects through space must be exclusively for peaceful purposes. Later, the concept of the peaceful uses of space was reflected in a great number of Resolutions, other UN documents, and State practice, and is now considered fundamental in space law. The question is what this concept practically means.

It is generally accepted that 'peaceful' does not mean 'non-military', rather it means 'non-aggressive'. This interpretation shares the fundamental principle of the UN Charter, which bans the threat or use of force, but allows force for self-defense and if sanctioned by the Security Council. Hence, any military space operation is lawful as long as it does not constitute a prohibited threat or use of force and does not otherwise violate international law, including space law.

---

[6] Marking the start of a new scientific and political era, the first artificial Earth satellite called Sputnik 1 was launched by the Soviet Union into an elliptical low Earth orbit on 4th October 1957. The transmitter batteries of Sputnik 1 were functioning for 21 days, while the satellite itself kept orbiting the Earth till 4th January 1958.

[7] *G.A. Res. 1148 (XII)*, U.N. GAOR, 12th sess. (1957), point 1(f): 'Urges that the States concerned <…>give priority to reaching a disarmament agreement which <…> will provide for the following: <…> f) The joint study of an inspection system designed to ensure that the sending of objects through outer space shall be exclusively for peaceful and scientific purposes.'

Limitations on Military Space Operations

Legally binding rules which impose specific limitations on military space activities are provided for in the Outer Space Treaty and the Moon Agreement. The Outer Space Treaty establishes a legal regime for both outer space and celestial bodies, which are treated somehow differently, while the Moon Agreement only covers celestial bodies.

Outer space from the perspective of the Outer Space Treaty

As regards outer space, there is a ban on nuclear weapons or any other weapons of mass destruction. States are prohibited from placing in orbit around the Earth any objects carrying nuclear weapons or any other kinds of weapons of mass destruction, installing such weapons on celestial bodies, or stationing such weapons in outer space in any other manner.[8]

This ban, however, does not address ballistic trajectories of objects carrying weapons of mass destruction. It means that the mere transit through space of a nuclear warhead, which can be launched from point to point on the Earth, is not prohibited by the Outer Space Treaty but governed by other applicable rules of international law.[9]

It worth saying that the UN space treaties do not define weapons of mass destruction. On the one hand, it is well-established that chemical and biological weapons are also considered weapons of mass destruction. However, due to the absence of permanent human life in near space, the consequences of the use of such weapons might be deferent from those on the Earth. On the other hand, due to the laws of physics, the use of some other types of weapons in space may have much more destructive consequences than on the Earth, where they are not considered weapons of mass destruction.

Finally, the Outer Space Treaty itself does not prohibit the placement of conventional weapons in space. But for some States limitations can exist.

---

[8] *Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, Including the Moon and Other Celestial Bodies*, 27th January 1967, 610 UNTS 205, art. VI, para. 1.

[9] For instance, we cannot say that a transit is possible for chemical and biological weapons which are banned.

No first placement of weapons in outer space

For instance, sometime ago,[10] Russia undertook a unilateral obligation not be the first to place any weapons in outer space, and since then has been encouraging other nations to follow the example. This political endeavor is supported by the UN General Assembly.[11]

As of today, 21 States have such a commitment.[12] For them, placement of conventional weapons in space is not permissible. At least, until any other State does it. It can actually occur quite soon as today we can see another trend as well. For instance, States are establishing space forces, decreasing vulnerability of their space assets and increasing their defense capabilities, including by the planned equipping them with weapons.[13]

Nuclear Weapons and Challenging Issues

As regards the use of nuclear weapons in space, which is prohibited by the Outer Space Treaty, consider this ban in the context of self-defense. The general question of the legality of the threat or use of nuclear weapons was earlier examined by the ICJ.[14] The Court could not *'conclude definitively whether the threat or use of nuclear weapons would be lawful or unlawful even in an extreme circumstance of self-defence, in which the very survival of a State would be at stake'*.

Another challenging issue is the use of nuclear weapons for planetary defense. It is argued that in such circumstances nuclear weapons may be the

---

[10] As far back as in 1983, the Soviet Union assumed an obligation not to be the first to station any kind of anti-satellite weapon in outer space. In 2004, Russia undertook a unilateral obligation not to be the first to place any weapon in outer space.

[11] *G.A. Res. 73/31*, U.N. GAOR, 73[d] sess. (2018), point 5: '*Encourages* all States, especially spacefaring nations, to consider the possibility of upholding, as appropriate, a political commitment not to be the first to place weapons in outer space.'

[12] Argentina, Armenia, Belarus, Bolivia, Brazil, Cuba, Ecuador, Guatemala, Indonesia, Kazakhstan, Kyrgyzstan, Nicaragua, Pakistan, Russian Federation, Sri Lanka, Suriname, Tajikistan, Uzbekistan, Uruguay, Venezuela, and Viet Nam.

[13] On the slide, several news items were shown to illustrate the trends in countries including the US, India, France and Japan.

[14] *Legality of the Threat or Use of Nuclear Weapons*, Advisory Opinion, ICJ Reports 1996, p. 226, at p. 266, para. 105: 'However, in view of the current state of international law, and of the elements of fact at its disposal, the Court cannot conclude definitively whether the threat or use of nuclear weapons would be lawful or unlawful even in an extreme circumstance of self-defence, in which the very survival of a State would be at stake.'

only option. I would say that the wrongfulness of the use of such weapons to destroy an asteroid approaching the Earth or a habitable space station could be precluded under the plea of necessity.[15]

Celestial Bodies from the perspective of the Outer Space Treaty

The legal regime of celestial bodies, in terms of their military use, is stricter than that of outer space. According to the wording of the Outer Space Treaty, the Moon and other celestial bodies must be used 'exclusively for peaceful purposes'.[16] Besides the ban on nuclear weapons and other weapons of mass destruction, the testing of any type of weapons is not allowed on celestial bodies. The establishment of military bases, installations and fortifications, and the conduct of any military maneuvers are also prohibited.

---

[15] The International Law Commission in the Commentaries to Draft Articles on Responsibility of States for Internationally Wrongful Acts (November 2001, Supplement No. 10 (A/56/10)), at p.81 referring to Affaire de l'indemnité russe, Russie, Turquie, 1912 (UNRIAA, vol. XI (Sales No. 61.V.4)), at p. 443: '… the obligation for a State to execute treaties may be weakened "if the very existence of the State is endangered, if observation of the international duty is ... self-destructive;" also see Articles on Responsibility of States for Internationally Wrongful Acts, art. 25: '1. Necessity may not be invoked by a State as a ground for precluding the wrongfulness of an act not in conformity with an international obligation of that State unless the act: (a) is the only way for the State to safeguard an essential interest against a grave and imminent peril; and (b) does not seriously impair an essential interest of the State or States towards which the obligation exists, or of the international community as a whole. 2. In any case, necessity may not be invoked by a State as a ground for precluding wrongfulness if: (a) the international obligation in question excludes the possibility of invoking necessity; or (b) the State has contributed to the situation of necessity;' it is also sometimes suggested that the wrongfulness of the use of nuclear weapons for the purposes of planetary defence may be precluded, if the situation does qualify as distress (see Articles on Responsibility of States for Internationally Wrongful Acts, art. 24: '1. The wrongfulness of an act of a State not in conformity with an international obligation of that State is precluded if the author of the act in question has no other reasonable way, in a situation of distress, of saving the author's life or the lives of other persons entrusted to the author's care. 2. Paragraph 1 does not apply if: (a) the situation of distress is due, either alone or in combination with other factors, to the conduct of the State invoking it; or (b) the act in question is likely to create a comparable or greater peril').

[16] *Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, Including the Moon and Other Celestial Bodies*, 27th January 1967, 610 UNTS 205, art. IV, para. 2. In the doctrine, two main views exist on the interpretation of the notion 'peaceful purposes'. The first one provides that celestial bodies are fully demilitarized, and any activity of military nature is prohibited on celestial bodies. The other viewpoint adopts a narrower interpretation stating that only those military activities are prohibited on celestial bodies which are directly listed in the second paragraph of Article IV of the Outer Space Treaty.

Celestial Bodies from the perspective of the Moon Agreement

The legal framework of military space activities on celestial bodies was further developed in the Moon Agreement. It introduces additional limitations, which are obligatory for 18 States.[17]

The Moon Agreement prohibits weapons of mass destruction not only on but also in celestial bodies. Another new limitation relates to orbits around, or other trajectory to or around, celestial bodies – they must also be free from weapons of mass destruction. By prohibiting the use of trajectories, the Moon Agreement seems to forbid gravity assistance from being used to redirect such weapons. As a consequence, objects carrying weapons of mass destruction must not transit along celestial bodies' orbits.

The Moon Agreement reiterates the prohibition of the threat or use of force, as specified in Article 2(4) of the UN Charter,[18] and prohibits any other hostile act or threat of hostile act. Neither the Moon Agreement, nor the *travaux préparatoires* provide details of what legal content was given by the drafters to the notion of a 'hostile act'. We can assume that there might be an act which is hostile in its nature but is less grave than the use of force, both being prohibited by the Moon Agreement.

Prior Consultations

Another set of rules, which comes close to the regulation of military operations in space, is a twofold mechanism of prior consultations.[19] On the one hand, such consultations must be undertaken; on the other hand, they may be requested.[20]

---

[17] Armenia, Australia, Austria, Belgium, Chile, Kazakhstan, Kuwait, Lebanon, Mexico, Morocco, Netherlands, Pakistan, Peru, Philippines, Saudi Arabia, Turkey, Uruguay, and Venezuela.

[18] For instance, when signing the Moon Agreement, France made a statement supporting exactly such an interpretation. See the Interpretative statement: 'France is of the view that the provisions of Article 3, paragraph 2, of the Agreement relating to the use or threat of force cannot be construed as anything other than a reaffirmation, for the purposes of the field of endeavor covered by the Agreement, of the principle of the prohibition of the threat or use of force, which States are obliged to observe in their international relations, as set forth in the United Nations Charter' (for the status of the Moon Agreement and reservations made hereto, see UN Treaty Collection, status webpage of the Moon Agreement).

[19] To date, consultations in accordance with this mechanism have been neither initiated nor requested.

[20] *Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, Including the Moon and Other Celestial Bodies*, 27th January 1967, 610 UNTS

This mechanism is triggered when a State has reason to believe that a planned space activity may cause potentially harmful interference to activities of other States. Though there is no definition of harmful interference in space law,[21] military operations may have an element of interference with space activities of other actors. For instance, space debris can be regarded as causing such interference. Hence, if a State plans a destructive military operation that creates space debris on orbits which are intensively used by other States, such a State is expected to undertake prior consultations.

It is important to say that space activities to which interference can be caused, must be peaceful. If not, the mechanism of prior consultations is not applicable.

You can also note that the wording leaves certain discretion to States. In deciding whether there is a reason to believe, a State should take into account all the available circumstances and assess them reasonably and impartially on a case by case basis.

Finally, the Outer Space Treaty neither obliges States to enter into proposed consultations, nor requires the States involved to reach a resolution of the issue, and no prior consent is necessary for a State to proceed with its planned space operation.

Responsibility for National Space Activities

What is also unique in space law, is the regime of international responsibility for national activities in outer space. If compared with the

205, Art. IX: '…If a State Party to the Treaty has reason to believe that an activity or experiment planned by it or its nationals in outer space, including the Moon and other celestial bodies, would cause potentially harmful interference with activities of other States Parties in the peaceful exploration and use of outer space, including the Moon and other celestial bodies, it shall undertake appropriate international consultations before proceeding with any such activity or experiment. A State Party to the Treaty which has reason to believe that an activity or experiment planned by another State Party in outer space, including the Moon and other celestial bodies, would cause potentially harmful interference with activities in the peaceful exploration and use of outer space, including the Moon and other celestial bodies, may request consultation concerning the activity or experiment.'

[21] In contrast, international telecommunications law does contain a definition of 'harmful interference'. See the Radio Regulations of the International Telecommunication Union. Edition 2016, Volume I, No. 1.169: 'harmful interference: Interference which endangers the functioning of a radionavigation service or of other safety services or seriously degrades, obstructs, or repeatedly interrupts a radiocommunication service operating in accordance with Radio Regulations'. Similar definition is specified in the Annex to the Constitution of the International Telecommunication Union, No. 1003.

customary law of State responsibility,[22] the threshold for the attribution of a conduct to the relevant State is lower. States are responsible not only for space activities of governmental agencies but also for the activities of non-governmental entities,[23] which include private companies and individuals.

This certainly applies to any space activity that is licensed by a State. It is also argued that (a) all space activities which are conducted on the territory of a State and (b) all space activities which are conducted by the State's national entities on any territory, are national space activities of that State for which it is responsible. Not only it is important when responsibility for internationally wrongful act is invoked, it may affect the determination of the parties of an armed conflict[24] and the application of neutrality law.[25]

International Liability

Now, let's discuss liability in space law.

It is a general rule, that a launching State is internationally liable for damage caused by its space object on the Earth, in air space, or in outer

---

[22] Under customary rules of State responsibility reflected in the Articles on Responsibility of States for Internationally Wrongful Acts (Art.s 4-11), for the conduct to be attributable to a State, it is necessary that certain circumstances are established and proved. The conduct of organs of a State, conduct of persons or entities exercising elements of governmental authority, conduct directed or controlled by a State, and conduct acknowledged and adopted by a State as its own can be attributed to a State in a particular situation.

[23] *Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, Including the Moon and Other Celestial Bodies*, 27th January 1967, 610 UNTS 205, Art. VI: 'States Parties to the Treaty shall bear international responsibility for national activities in outer space, including the Moon and other celestial bodies, whether such activities are carried on by governmental agencies or by non-governmental entities, and for assuring that national activities are carried out in conformity with the provisions set forth in the present Treaty. The activities of non-governmental entities in outer space, including the Moon and other celestial bodies, shall require authorization and continuing supervision by the appropriate State Party to the Treaty.'

[24] For instance, would a State be regarded a party to an armed conflict if its national non-governmental entity, which is duly licensed (scenario 1) or which acts with no license (scenario 2), enters into an armed conflict due to an activity in outer space, since such an activity, according to international space law, would be (or could be) attributed to that State as that State's national activities in outer space?

[25] For instance, during an armed conflict, it would be necessary for neutral States to terminate services that are not neutral, including those provided by non-governmental entities.

space.[26] This rule will be suspended between the belligerents and will not be applicable to armed conflicts. Still, it is relevant to military space operations in peacetime.

It is important here, that liability can only be invoked if damage is caused by a space object, for example, as a result of a physical collision. If damage is caused not by a space object, for instance, by the use of the radio-frequency spectrum, it will not be covered by the rules of liability.

Registration

Another set of rules which is relevant to space objects, is the registration regime.

The Registration Convention requires that the launching State registers its space objects and submits information to the UN Secretary-General.[27] The submission of information by States which are not bound by the

[26] *Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, Including the Moon and Other Celestial Bodies*, 27th January 1967, 610 UNTS 205, art. VII: 'Each State Party to the Treaty that launches or procures the launching of an object into outer space, including the Moon and other celestial bodies, and each State Party from whose territory or facility an object is launched, is internationally liable for damage to another State Party to the Treaty or to its natural or juridical persons by such object or its component parts on the Earth, in air space or in outer space, including the Moon and other celestial bodies.' The provision was further developed in the *Convention on International Liability for Damage Caused by Space Objects*, 29th March 1972, 961 UNTS 187, which is usually regarded as *lex specialis* to Article VII. Article II of the Convention provides for regime of absolute liability: 'A launching State shall be absolutely liable to pay compensation for damage caused by its space object on the surface of the Earth or to aircraft in flight,' while its Article III – for fault-based liability: 'In the event of damage being caused elsewhere than on the surface of the Earth to a space object of one launching State or to persons or property on board such a space object by a space object of another launching State, the latter shall be liable only if the damage is due to its fault or the fault of persons for whom it is responsible.'

[27] *Convention on Registration of Objects Launched into Outer Space*, 14th January 1975, 1023 UNTS 15, Art. II para. 1: 'When a space object is launched into Earth orbit or beyond, the launching State shall register the space object by means of an entry in an appropriate registry which it shall maintain. Each launching State shall inform the Secretary-General of the United Nations of the establishment of such a registry;' Art. IV para. 1: 'Each State of registry shall furnish to the Secretary-General of the United Nations, as soon as practicable, the following information concerning each space object carried on its registry: (a) Name of launching State or States; (b) An appropriate designator of the space object or its registration number; (c) Date and territory or location of launch; (d) Basic orbital parameters, including: (i) Nodal period; (ii) Inclination; (iii) Apogee; (iv) Perigee; (e) General function of the space object.'

Convention can be, and actually is, performed on a voluntary basis, in accordance with the Resolution of the UN General Assembly.[28]

The registration regime covers all space objects, including dual-use and military ones. Today, States are registering such satellites, however, it remains States' discretion how their general function is described.

For instance, the Athena-Fidus satellite, which is known to serve French defense, is described by France as 'telecommunications satellite' with no reference to its use for military space activities. To compare, Eutelsat 3B, a purely commercial satellite, is given the same description by France. In these cases, the UN Register does not help to determine whether these space objects are military or non-military. Alongside these examples, there are examples when the military nature of space objects is disclosed. Please see for yourself these examples on this slide[29] that States are given flexibility to determine what information is submitted to the UN.

In doing so, States should keep in mind that the Register is a primary source of information on space objects and is public. It can be used by attack planners, when targeting, in order to verify that the potential target is a legitimate military objective and not a civilian object. Therefore, if the general function of a purely military satellite is, during an armed conflict, intentionally deceptively described as civilian, such an act may be viewed as perfidy.

What is also important from the military perspective, every space object requires registration. For instance, when swarms of nano-satellites are co-launched in space to orbit a strategic satellite, thereby inspecting or protecting it, each of them must be registered.

The good news is that humans in space are not required to be registered.

Rescue and Return

Under international space law, astronauts are considered the 'envoys of mankind'. As such, they must be rendered all possible assistance in the

[28] *G.A. Res. 1721B(XVI),* U.N. GAOR, 16th sess. (1961), point 1: '*Calls upon* States launching objects into orbit or beyond to furnish information promptly to the Committee on the Peaceful Uses of Outer Space, through the Secretary-General, for the registration of launchings.'

[29] For instance, CSO 1 which is described by France as a 'defense satellite'; SICRAL 2 which is identified by Italy as 'military telecommunications satellite'; all the Russian Cosmos series satellites are generally identified as 'intended for assignments on behalf of the Ministry of Defense of the Russian Federation'; the SKYNET 5D satellite which, as notified by the United Kingdom, 'provides secure military communications capability to British Armed Forces and friendly nations'.

event of distress and be returned to their States. More so, there is a duty for astronauts to render assistance to each other.

Here, it is important to say that astronauts, even those who participate in civil space programs, are often members of the military. So, a question arises whether all astronauts, including those involved in military space operations in peacetime, are entitled to the same level of protection. Neither the UN space treaties, nor State practice[30] distinguish between military and non-military astronauts. Hence, taking into account the 'sentiments of humanity', the protection in peacetime seems to equally apply to all astronauts.

However, it would seem reasonable to assume that the outbreak of an armed conflict could constitute a 'fundamental change in circumstances',[31] which could change an astronaut's status from that of an 'envoy of mankind' to that of a 'combatant'. Even though IHL allows for the targeting of 'all members of the armed forces, whether or not they are actually engaged in combat',[32] the engagement of astronauts in military space activities supporting combat operations should be assessed on a case by case basis.[33]

---

[30] For instance, Ambassador Arthur Goldberg, when reporting on the Outer Space Treaty drafting and negotiations to the US Senate Committee on Foreign Relations, expressly stated that agreement was reached during negotiations on the point that the protection shall be applied unconditionally to all astronauts, including military persons. See Statement by Ambassador Goldberg, 'Hearings Before the Committee on Foreign Relations, United States Senate', 90th Congress, 1st Session, 1967.

[31] *Vienna Convention on the Law of Treaties*, 23rd May 1969, 1155 U.N.T.S. 331, Art. 62: 'A fundamental change of circumstances which has occurred with regard to those existing at the time of the conclusion of a treaty, and which was not foreseen by the parties, may not be invoked as a ground for terminating or withdrawing from the treaty unless: (a) the existence of those circumstances constituted an essential basis of the consent of the parties to be bound by the treaty; and (b) the effect of the change is radically to transform the extent of obligations still to be performed under the treaty. 2. A fundamental change of circumstances may not be invoked as a ground for terminating or withdrawing from a treaty: (a) if the treaty establishes a boundary; or (b) if the fundamental change is the result of a breach by the party invoking it either of an obligation under the treaty or of any other international obligation owed to any other party to the treaty. 3. If, under the foregoing paragraphs, a party may invoke a fundamental change of circumstances as a ground for terminating or withdrawing from a treaty it may also invoke the change as a ground for suspending the operation of the treaty.'

[32] Yoram Dinstein, *The Conduct of Hostilities under the Law of International Armed Conflict* (2004), p. 94. See also Michael N. Schmitt, *State-Sponsored Assassination in International and Domestic Law*, 17 Yale J. Int'l L. (1992), p. 674: 'Second, lawful targeting in wartime has never required that the individual actually be engaged in combat. Rather, it depends on combatant status'.

[33] The mere fact that there is an enemy's astronaut on the International Space Station should not be enough to target him or her. Since in case of outbreak of an armed conflict

Rescue and return obligations are also set forth with regard to space objects, however, with a lower degree of dedication on the part of States.[34] The question of space objects in an armed conflict seems to be an easier one. While the Rescue Agreement is suspended between belligerents, the enemy's space objects can be captured and destroyed, provided that other applicable rules of international law are complied with.

Fundamental Principles of Space Law

When conducting military space operations, States should also take into consideration other fundamental principles of international space law which can be found in the Outer Space Treaty.[35] Some of them are considered to be customary in nature, but the scarcity of State practice makes it legally

---

international space law and international humanitarian law will both constitute *legi speciali* regarding the status of astronauts, such case leads to the conflict of laws. A customary rule known as the 'Martens Clause' provides, *inter alia,* that in situations which are not covered by specific provisions of international law, conduct in the armed conflict shall be governed by the principle of humanity. It seems reasonable to assume that the principle of humanity analogously can apply in the situation of conflict of laws.

[34] Compare the wording of the Rescue Agreement in the context of space objects, which is 'take such steps as it finds practicable to recover the object or component parts' (Art. 5, para. 2), with the wording of the Rescue Agreement concerning astronauts, which is 'immediately take all possible steps to rescue them and render them all necessary assistance' (Art. 2).

[35] Above all is the freedom of use of outer space enshrined in Art. I of the Outer Space Treaty. It is exactly this principle which makes it possible to launch satellites for telecommunications, broadcasting, and remote sensing. What is important, it can be done without seeking a permission from the State which territory is overflown by the satellite. Even though such a State might not be happy with a foreign satellite's imaging its territory from space, it is limited in the choice of measures that can be applied lawfully. For instance, it can conceal its critical infrastructure, however, an intentional dazzling or blinding of a foreign satellite to keep it from viewing a specific area can violate the other State's right to use outer space freely.

Closely related to the freedom of exploration and use, is the principle of non-appropriation of outer space provided for by Art. II of the Outer Space Treaty.

Another principle, which says that the exploration and use of outer space shall be carried out for the benefit and in the interests of all countries and shall be the province of all humankind (Art. I of the Outer Space Treaty), obliges States to look beyond their respective purely national interests when conducting space activities.

This principle is given substantive effect in other provisions of the Outer Space Treaty, which require that in the exploration and use of outer space States shall be guided by the principle of cooperation and mutual assistance and shall conduct all their activities in outer space with due regard to the corresponding interests of other States (Art. IX of the Outer Space Treaty).

complicated to correctly apply these principles to military space operations in peacetime. Their application to military space operations in times of hostilities is even more challenging. In this regard, let me tell you about some recent developments.

Recent Developments

At least two international projects are currently being implemented, which are aimed at objectively articulating and clarifying international law applicable to military activities in outer space. The first one is called MILAMOS[36] where an international group of experts is drafting the Manual on International Law Applicable to Military Uses of Outer Space and I am honored to participate in this project as Core Expert and Associate Editor. The other project is called Woomera[37] where an international group of experts is working on the Manual on the International Law of Military Space Operations and I am happy to see honored experts from this project in the room.

Here, in the city of Sanremo, where the Manual on International Law Applicable to Armed Conflicts at Sea has originated, it is needless to explain the importance of such manuals for the promotion of the rule of law and for ensuring its common understanding. Conflicts in space are not inevitable and international cooperation can help avoid tough scenarios and protect the unique space domain, so it remains available for the benefit of the current and future generations in all States. That is, for sure, our common desire.

This brings me to the end of my presentation and, hopefully, opens promising discussions. Thank you all for listening.

---

[36] The MILAMOS Project is aimed at the creation of the manual articulating and clarifying existing international law applicable to military uses of outer space in time of peace, including challenges to peace. The project is carried out under the auspices of McGill Centre for Research in Air and Space Law, in cooperation with partner institutions, with the expected date of release in 2020. For details, please visit https://www.mcgill.ca/milamos/.

[37] The Woomera Project is aimed at the creation of the manual articulating and clarifying existing international law applicable to military space operations being therefore focused on the time of an armed conflict. The project is carried out under the aegis of the University of Adelaide, the University of Exeter, the University of Nebraska, and the University of New South Wales – Canberra with the expected date of release in 2021. For details, please visit https://law.adelaide.edu.au/woomera/.

How does IHL apply in outer space and which challenges exist for applying existing rules in outer space?

Liang JIE
Associate Professor, National Defence University PLA China

Since the 1950s, with the development and successful launch of different spacecraft such as satellites, humankind began to enter the vast outer space. In ancient Chinese art of military, it says if you control a higher terrain, you can easily win the war. The space is extensive and boundless, and it can provide us a wide field of vision. Many countries explored outer space for military purposes from the start.

Although the Charter of the United Nations regulates, 'All Members shall settle their international disputes by peaceful means in such a manner that international peace and security and justice, are not endangered.' The Charter also requires all Members refrain in their international relations from the threat or use of force against the territorial integrity or political independence of any state, or in any other manner inconsistent with the purposes of the United Nations. At the same time, the Charter recognizes that all states have the inherent right of individual or collective self-defense.

Although in the Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, including the Moon and Other Celestial Bodies, regulates 'States Parties to the Treaty shall carry on activities in the exploration and use of outer space, including the Moon and other celestial bodies, in accordance with international law, including the Charter of the United Nations, in the interest of maintaining international peace and security and promoting international cooperation and understanding.' And the Treaty also regulates, 'The Moon and other celestial bodies shall be used by all States Parties to the Treaty exclusively for peaceful purposes.'

There are different understandings about 'peaceful purposes'. Some hold the opinion that 'peaceful purposes' means 'non-military'. Once the exploration and use of outer space are related to military issues, such as carrying out military investigation by satellites or placing weapons in outer space, these activities are not consistent with 'peaceful purposes'. Others think that the opposite meaning of 'peaceful' is 'aggressive'. So long as the

outer space is not used for aggressive purposes, even if it's used to provide military services or to place weapons, the use is lawful and permissible.

We can conclude from states' practice that militarization is an inevitable trend in outer space exploration and use. Undoubtedly, outer-space warfare will be an important part of future armed conflict.

Outer-space warfare is the military confrontation between states. It not only includes military attack and defense happened in outer space and actions taken in outer space with its damage effects occurred in air space or on the earth, but it also includes actions taken in air space or on the earth aiming at destroying or invalidating outer-space systems.

Parties to outer-space warfare should abide by international humanitarian law. International humanitarian law is composed of a set of rules which limits the means and method of warfare and protects the victims of armed conflict. We know that law usually lags behind reality. There are no specific legal rules applicable to outer-space warfare up to now. However, the fundamental principles of international humanitarian law should apply for the fundamental principles embody the essence and core value of international humanitarian law. They are the basic criterion for choosing means and method of warfare.

Now let's talk about the legal challenges to the application of fundamental principles of international humanitarian law to outer-space warfare.

The principle of distinction

The principle of distinction requires that the Parties to the conflict shall at all times distinguish between the civilian population and combatants and between civilian objects and military objectives and accordingly shall direct their operations only against military objectives. Although in the Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, including the Moon and Other Celestial Bodies, it regulates that: 'States Parties to the Treaty shall regard astronauts as envoys of mankind in outer space'. Astronauts operating military spacecraft in outer-space warfare undoubtedly are combatants and can be legally targeted and attacked. In traditional armed conflict the symbols identifying combatants are their military uniforms. However, in outer-space warfare, spacecraft are tightly sealed and move at high speed, the combatants inside can't be seen directly.

So, the criterion to choose military targets is the character of the spacecraft. That means it depends on whether the spacecraft is military or civilian. However, in practice most countries develop their space industry by the way of civil-military integration. Many spacecrafts can be used for either military or civilian purposes. It's hard to distinguish military objectives from civilian objects promptly. Besides armed forces, according to Additional Protocol I, Article 52, military objectives are limited to those objects which by their nature, location, purpose or use make an effective contribution to military action and whose total or partial destruction, capture or neutralization, in the circumstances ruling at the time, offers a definite military advantage.

In view of "nature" standard, military satellites are military objectives undoubtedly, even if they also have civil functions. But if ordinary satellites such as commercial communication satellites, navigation satellites, remote sensing satellites and meteorological satellites provide services to armed forces, they can be legally attacked only when the evidence is conclusive.

The "purpose" or "use" standard is more difficult to judge. When a spacecraft only has potential military functions such as its design parameters that meet the military standards, it cannot be attacked as a military objective. However, when there is convincing evidence showing that a civil spacecraft has an intention to take direct part in hostilities, for example, a satellite registered for civil use suddenly changes its orbit and approaches the enemy military spacecraft, this obvious intention to attack will be the legal evidence that it has become a legal target. The difficulty in practice is that it's hard to identify the attack intention. Probably only after being attacked, can a party to the conflict make a judgment.

The 'location' standard may include outer space orbit in the scope of military objectives. In land warfare, if the enemy has occupied a piece of land which provides the enemy a geographic advantage, this piece of land can be attacked as a legal target. For the same reason, in outer-space warfare, if a particular outer space orbit can be used by the enemy to observe military actions, transfer military information or conduct military operations, a party to the conflict has the right to prevent hostile parties from reaching the orbit. Parties to the conflict can even cause an explosion in the particular area. However, this will lead to space debris problems. For according to international humanitarian law, parties to the conflict have the obligation to protect the natural environment. Article 55 to Additional Protocol I requires parties to the conflict 'to protect the natural environment against widespread, long-term and severe damage'.

Besides the military objectives deployed in outer space, the identity of the personnel who work on the ground is also difficult to identify. According to international humanitarian law, civilians shall enjoy general protection against dangers arising from military operations, unless they take a direct part in hostilities. “Direct” participation means acts of war which by their nature or purpose are likely to cause actual harm to the personnel and equipment of the enemy armed forces. In outer-space warfare, when technicians of Commercial satellite co., Ltd. maintain or repair dual-use satellites, it’s hard to decide whether they can be legally attacked.

The principle of proportionality

The principle of proportionality requires the Parties to the conflict should refrain from launching any attack which may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive in relation to the concrete and direct military advantage anticipated. The foregoing principle of distinction is about how to choose legal targets, and the principle of proportionality goes a step further, it’s aimed at avoiding or minimizing incidental damages when attacking military objectives.

Military confrontation in outer space can use either soft or hard means of warfare. The soft means is also called interference type attack. It will prevent the enemy satellites from receiving signals and then prevent them from working properly. People may think that this kind of means of warfare doesn’t destroy the satellites directly and will cause less incidental loss and damage. So, they think that the soft attack is more humane and more in line with the principle of proportionality.

Actually, however, military confrontation in outer space involves huge combat systems. Many military facilities depend heavily upon satellite positioning systems. Once the satellite positioning systems are interfered with, the weapons will not be able to target accurately, they may attack non-military objectives and violate the principle of proportionality.

For example, State A attacks State B’s satellite and render its system ineffective. State B’s satellite sends error signals to its space shuttle. The space shuttle drops bomb in accordance with the wrong instructions, and the bomb deviated from its original target and exploded, damages civilian objects. Such an attack is likely to violate the principle of proportionality.

The hard means of warfare attack targets directly and destroys them physically. This kind of means usually uses directed energy weapons such as missiles. It is the most effective means of warfare to control outer space and to defeat enemies. However, the space debris it causes will move rapidly in outer space, and can collide with spacecrafts, no matter if the spacecrafts are military or civilian, and no matter who they belong to, they would probably be destroyed completely. At the same time, our daily life depends upon space technology more and more. It reflects in many aspects such as health care, finance and transport. Attacks to satellites may cause incidental loss of civilian life and damage civilian objects severely.

When parties of the conflict attack military objectives on the ground from outer space, even though space-based weapons usually are accurate and intelligent, they have great lethality and will cause severe damages on a large scale. It will likely be going against the principle of proportionality.

In these cases, for a reasonable commander it is hard to judge whether his order is in line with the principle of proportionality.

Moreover, Article 58 of Additional Protocol I requires parties to the conflict to take necessary precautions to protect the civilian population, individual civilians and civilian objects under their control against the dangers resulting from military operations. In outer-space warfare, space forces often provide information support to military operations and they are usually attacked by surprise. So, it's not realistic to take precautions such as giving advance notifications or evacuating civilians before attacks.

Conclusion

The foregoing analysis shows that there are huge legal challenges for international humanitarian law when it applies to outer-space warfare. That means existing international humanitarian law is not clear enough to guide and regulate outer-space warfare. It should be strengthened.

The international community does not have a unified legislature, and international legal rules regulating outer-space warfare can only be formulated by sovereign States reaching agreements. However, the ability of military exploration and use of outer space varies greatly from country to country. Based on different realistic needs, countries have different interpretations of the principles of international humanitarian law. It's difficult to come to an agreement.

Considering this international social background, we can use the ‘soft law’ form. For example, international organizations or international conferences can reach resolutions or declarations, and academic institutions can offer proposals. Although these documents don’t have legal effects and cannot bind the parties of an armed conflict, they are helpful to the final formation of international humanitarian rules.

In short, outer-space warfare is a new type. Compared with traditional armed conflict, the means and methods it uses are different. Existing international humanitarian law was developed from traditional warfare, and it encounters some difficulties when applying to outer-space warfare. We should draw attention to this phenomenon, study the legal challenges intensively, communicate and co-operate closely, and try our best to establish new rules which can reflect the common interests of the international community.

Military implications of the use of outer space: a European perspective

Jérémie AYADI
Captain, Legal Advisor, French Joint Space Command,
French Ministry of Defence

European countries, individually or collectively, hardly approached space from an exclusively military angle. For a long time, Europe has restricted itself to purely civilian programs of a scientific character as the European Space Agency (ESA) was designed for exclusively peaceful purposes, excluding any development of specific military space assets. However, it did not preclude any military program developed at a national level, whether individually or in cooperation.

Moreover, space technologies do not easily respect such a *summa division*. The main example of this reality can be found in the very beginning of the space conquest. Indeed, the launch of Sputnik in 1957 did not only demonstrate to the whole world Soviet engineers could master orbit injections before the United States: it assured the capability of the Soviet Union to launch a ballistic missile with a 6 000 km range. Since the beginning, space technologies could be considered as dual by nature.

This duality can be observed in most, if not all elements constituting an independent or autonomous space power, such as access to space, telecommunication, remote-sensing and meteorological satellites, global navigation satellite system, electronic intelligence satellites or space situational awareness. In Europe, these types of capabilities were developed often sequentially over 55 years at different levels, benefitting each other thanks to national efforts, through intergovernmental cooperation but also under a supranational organization, EU, which finds diplomatic and military tools in primarily civilian assets before developing proper dual-use capabilities.

The European access to outer space is tightly linked to the European Launcher Development Organisation (ELDO), the European Space Agency (ESA) but also to the French effort of having access to a launcher. When some European States created ELDO in 1963, their main goal was to develop a regional launcher system, *Europa*, as a prelude to an autonomous space program of scientific nature. Although *Europa* definitely failed in 1972, France advocated in favor of another launcher in 1973 under management of the CNES. It became known as *Ariane* and is still the

prominent launcher of ESA founded in 1975 and benefited France in the development of its ballistic nuclear missiles after the end of the "precious stones" program. Indeed, the need for an autonomous launching system was key for the credibility of the French *force de dissuasion* as freedom of access to space is a prerequisite of any space strategy.

Telecommunication satellites or SATCOM are the most commonly used satellites as they are key for modern forces relying on encrypted and quick coordination. If the European States worked together on the OTS program from ESRO and then ESA to launch their first SATCOM in 1978, the first Western European country to develop a military one was the United Kingdom thanks to the Skynet programme, the first being launched in 1969, and which is now running its 6th series.

France started developing its own dual-use satellites in 1980 with the SYRACUSE 1 and 2 programs with several launches from 1984 to 1996. In 2005, the first satellite of the military SYRACUSE 3 series was put into orbit. Apart from the United Kingdom, Italy has also specialized in military SATCOM, with the SICRAL 1 series launched in 2001 and 2009 but also two satellites jointly developed with France: the dual-use Athena-Fidus and the military SICRAL 2. Germany has its own military satellites known as SATCOMBw-1 and SATCOMBw-2 launched in 2009 and 2010. The strategic and critical nature of a secured national SATCOM constellation may explain why there was less cooperation in this very domain.

However, the EU decided in 2013 to undertake a dual-use program called the GOVSATCOM initiative which consists in a dedicated platform of pooling and sharing of governmental, both civilian and military, and commercial satellites to allow EU members to access to secure or available telecommunication assets.

Remote-sensing or Earth observation satellites benefited from civilian and military demands: if the launchers could deliver a conventional or a nuclear device onto a target, a remote-sensing satellite can gather imagery of these systems from the highest point and without any legal constraint: as a "national technical means", it serves a strategical purpose in arms verifications and readiness of the military apparatus.

In 1977, EUMETSAT was the initial European organization to operate a weather satellite, the METEOSAT series developed by ESA. The need for military imagery continued to increase due to the Cold War era and France launched in 1978 its first military Earth's observation program, abandoned in 1982. France, Belgium and Sweden in 1978 started to develop dual remote-sensing satellites with the SPOT program, benefitting from the French military one, the first being launched in 1986. The French HELIOS

I military program started in 1985 with the participation of Italy and Spain, for the highest political making authorities rather than for tactical purposes. In 1998, another French military program, HELIOS II, was launched with the participation of Italy and Spain, Belgium and Greece, and two optical remote-sensing satellites were launched in 2005 and in 2009.

The need for more precise weather data led the ESA members to undertake the METEOSAT program for low Earth's observation scientific satellites: the ERS series from 1991 to 1995, with a *synthetic aperture radar*, which would become a specialty of both Italy and Germany. However, ESA was not the only major European organization interested in outer space: the European Union (EU) and its Commission. Indeed, the negotiations on the 1997 Kyoto Protocols highlighted the need for global environment data, leading to the EU's Global Monitoring for Environment and Security (GMES)/Copernicus program developed with ESA after the publication of the Baveno Manifesto in 1998. By 2019, seven Sentinels were being exploited.

However, from the fight against global warming, the Copernicus program had unexpected developments in security, as many programs and techniques used to monitor the environment also had security application. For example, the GMES services for Management of Operations, Situation Awareness and Intelligence for regional Crises or G-MOSAIC which is a 2009 project founded under the Common Security and Defence Policy, a main component of the EU's Common Foreign and Security Policy. Its missions occasionally consist in monitoring EU borders and areas of conflicts of interest to the EU. Before Copernicus, EU Satellite Centre (SATCEN) was created in 1992 in Torrejon in order to exploit imagery, supplied by EU members with classification restrictions while buying some of its data from private companies.

Regarding the military capabilities of European States, the necessary reliance upon foreign assets, more particularly during the Balkan wars, placed many allies in a situation of dependency, whether for political decision-making or for the conduct of terrestrial operations.

As a consequence, several countries decided to cooperate in other programs. In 2001, France and Italy signed the Torino Agreement and started the ORFEO program. Previously, and having participated in this METEOSAT program, Italy developed technologies for its own dual remote-sensing radar satellites known as COSMO-SKYMED, launched from 2007 to 2010. Under the French-Italian Agreement, France provided a dual optical component called *Pléiades* while Italy delivered its dual radar satellites as the radar component. For the same reasons, Germany

undertook its own military radar system called SAR-Lupe and, in 2002 signed the Schwerin Agreement with France, leading to the exchange of imageries between the German SAR-Lupe and the newest generation of French military optical satellites called HELIOS II.

In 2010, France started the CSO program to replace its HELIOS II satellites and in 2015, the Multinational Space-based Imaging System for Surveillance, Reconnaissance and Observation or MUSIS engaged France, Italy, Belgium, Germany, Greece, and Spain (Sweden joined later) to develop a military interoperable system able to run the military or dual-use French CSO, German SARah, Italian CSG and Spanish Ingenio.

However, the Schwerin Agreement and the subsequent specialization of both countries would be questioned in 2017 when Germany decided to finance an intelligence optical remote-sensing constellation called Georg, constructed by the German OHB company, while developing the SARah, successor to SAR-Lupe.

Space-based Positioning, Navigation and Timing assets, or Global Navigation Satellite System (GNSS) are of utmost importance, yet in an indirect way, for economic growth in developed countries. They definitely improve synchronization of banking transfers, secure the air traffic while optimizing travel distances and costs. They offer civilian ships an immediate positioning in case of emergency and they even help in precision agriculture and autonomous use of harvesters. Finally, GNSS enables autonomous spacecraft navigation, without guidance from Earth.

Due to its legal competences in economy, the EU finances its own GNSS program, reducing the dependency of its member States in the American GNSS: the GPS. The European Commission's initial proposal for the Galileo program was drawn up in 1999 by its transport and energy directorate. The European Commission delegated to the ESA the technical management of the development phase of the program on the condition that it obeyed EU management rules, which are different from those of the ESA and which exclude the guarantee of geographical benefits. Since the end of 2016, some of Galileo's services have been operational, including the free Open Service.

In the military field, GNSS are key for the coordination and positioning of forces all around the world. Moreover, they are paramount for modern targeting and the use of precise guided ammunitions, contributing to the proportionality principle of the law of armed conflicts. For Galileo, a Public Regulated Service, reserved for EU Member States' governmental services, should be available from the beginning of 2020. While it is not primarily dedicated to weaponry, but more to public order matters, France decided to

finance a program to use Galileo PRS for military purposes and to equip the armed forces with jam and spoofing-resistant receivers. The program called OMEGA will propose receivers to equip high-value platforms, effectors and ammunition in 2020, able to use both Galileo and GPS constellations.

Space surveillance and tracking (SST) are also one of the main components of spacefaring nations. Due to a combination of telescopes, lasers and radars, it is possible to detect and track a variety of satellites orbiting a LEO and GEO and compute them into a catalogue. This is a key factor to reduce the risk of satellites colliding with each other. It will also help, in a more and more congested outer space, to better determine windows of opportunities when launchings are programmed, in order to avoid a collision between the rocket and debris. But these capabilities can also be used in order to detect satellites not being registered in UNOOSA and having a less friendly behavior. By adding data from intelligence agencies to the catalogue, it is possible to determine monitor specific satellites for military protection of territories: the SST changes its nature into space situational awareness or SSA.

It was not by chance that European militarized powers developed their SSA capabilities, such as the French GRAVES and the German TIRA radars, but also the other dual-use detection and tracking networks were developed in European countries: as modern forces and countries are relying more and more on outer space, the protection of their space assets, and first of all the detection of any risk or threat, is a an important step toward a space defense strategy to protect space assets. Indeed, it is important to highlight that any attempt made on space assets could be detected, attributed and characterized, and accompany itself of diplomatic or forcible consequences.

Aware of the importance of SST in the preservation of both satellites and space environment, the EU decided in 2014 and 2016 to promote a consortium of Member States (EUSST) to share their catalogues, helping France, Germany, Italy, Spain, United Kingdom, Poland, Portugal and Romania to finance their SST assets, proportionally to the portion used for EUSST.

Scientific cooperation, especially thanks to ESA organization, triggered a virtuous circle where European national States promoted technologies to the point where they adapted these inventions to their national need and industrial apparatus. The specialization they developed and the proximity of some of their strategical objectives led to a kind of Ricardian

“comparative advantage”: some of them cooperated in an exchange of capacities.

However, the dual nature of space technologies incited the European Union to also develop its own space policy and programs to enhance the industrial fabric of the Old Continent while slowly increasing its interest in security, including military, applications of outer space.

The European amalgam of both civilian and military developments and objectives of outer space can be observed in the European Union Draft Code of Conduct for Outer Space Activities, released in 2008 in COPUOS. Indeed, the Draft Code, approved by the Council of the European Union, intends to deal with both safety and security in outer space, emphasizing the implementation of national policies to prevent accidents and collisions in outer space, refraining from the intentional creation of debris in outer space but also underlining the inherent right of individual or collective self-defense in accordance with the United Nations Charter. The importance of the freedom of access to, exploration and use of outer space and exploitation of space objects for peaceful purposes without interference is strongly affirmed but the complete demilitarization of outer space is not asserted as it would prevent the effectiveness of self-defense in outer space.

As such, the recent publication of the French Defence Space Strategy on 25th July 2019, and the objective to develop and launch future active defense capabilities in outer space, is another step in Europe towards the recognition of outer space as a conflictual domain due to its indubitable importance in the sustainment of our complex economies and space operations support to terrestrial forces.

It is now, with the ever-increasing pervasiveness of outer space in military affairs and economic development that some European States are reforming their governance and developing doctrines entirely dedicated to militarization and even weaponization of outer space. It is one of the logical consequences of outer space technologies and environment, as dual use by nature.

VI. New technology and urban warfare: more precise or more destructive?

Guerre urbaine en 2035 : à quelles réalités s'attendre ?

Xavier LABARRIÈRE
Colonel, Conseiller juridique en droit opérationnel, Quartier Général du Commandement suprême allié Transformation de l'OTAN

Je remercie d'abord l'Institut de Droit International Humanitaire de Sanremo pour son invitation. Je suis très honoré de pouvoir participer à cette table ronde en présence d'une audience aussi prestigieuse et exigeante et également enchanté de pouvoir m'exprimer dans l'autre langue de l'OTAN.

Je suis actuellement affecté à l'OTAN en qualité de Conseiller Juridique opérationnel au HQ ACT de Norfolk aux États-Unis. C'est le commandement stratégique de l'alliance en charge de la transformation. Il s'agit ici de sa préparation aux conflictualités de demain et le maintien de son avantage décisif dans un monde en redistribution.

Il m'a été demandé de m'exprimer sur le thème de la guerre dans les environnements urbains et en réalité auxquelles on peut raisonnablement s'attendre en 2035.

Alors, avant de commencer je souhaite préciser le champ de mon intervention qui se limitera volontairement à l'OTAN et au point de vue du commandement opérationnel et stratégique. Cette limitation est assez importante et je tenais à la préciser dès le départ.

Il y a selon moi deux sous questions dans le thème que je vais aborder aujourd'hui : la première est une question de prospectives et de stratégies au sens large, et la seconde, en relation directe avec l'objet de cette table ronde, pourrait être formulée dans la manière suivante : quelles sont les problématiques juridiques qui découlent des réponses, ou des orientations, qui pourraient être fixées dans le premier questionnement ?

Le but aujourd'hui est de ne pas consacrer trop de temps à la première partie, mais de présenter succinctement les méthodes et les orientations retenues par l'OTAN dans cet exercice de prospectives.

Le sujet de la guerre en environnement urbain n'est pas un sujet nouveau pour l'OTAN ; il a fait l'objet des travaux importants qui ont débouchés sur l'adoption d'un concept sur les opérations interarmées en environnement urbain le 27 Novembre 2018. Je vais en développer les principaux points saillants un peu plus loin. Mais s'agissant de prospectives stratégiques, et nous sommes ici au cœur du métier du HQ ACT, il s'agit

bien de se pencher sur les perspectives réalistes auxquelles on peut s'attendre en 2035.

La transformation militaire à long terme (LTMT, *long term military transformation*) de l'OTAN vise à établir une feuille de route pour la transformation des forces de l'OTAN en 2035. Cette transformation militaire à long terme est un processus basé sur un cycle de quatre ans avec trois temps. Premier temps, c'est l'édition du « Strategic *Foresight Analysis* (SFA) » ; second temps, c'est l'édition du « *Framework for Future Alliance Operations* (FFAO) », ces deux documents visant à orienter les décisions relatives aux cycles capacitaires, le NDPP, « *NATO Defence Planning Process* ». Ces deux documents, SFA et FFAO, vont également orienter les travaux des doctrines et les programmes d'éducation et d'entraînement.

Alors, on va rapidement balayer le SFA et le FFAO pour voir ce qu'ils contiennent dans le domaine de la guerre en environnement urbain. Le dernier SFA date d'octobre 2017. Le SFA est la présentation d'une vision partagée des enjeux stratégiques vus par les 29 membres de l'OTAN. On peut le résumer dans la manière suivante : c'est le « *je comprends le monde* ». Alors, c'est une compréhension qui s'appuie vraiment sur une approche globale à la fois interministérielle, académique et qui intègre les contributions de nombreux spécialistes de la société civile. Je précise également que SFA et FFAO sont deux documents qui sont disponibles en source ouverte. Dans les 20 tendances retenues par le SFA on va trouver que l'urbanisation croissante est un des aspects importants. Cette tendance va se retrouver déclinée en cinq implications :

- Rapidement, l'urbanisation croissante va conduire à une compétition augmentée sur les ressources ;
- Seconde implication, l'urbanisation va conduire à ce que la propriété et le contrôle des infrastructures critiques soient contestés ;
- Troisième implication, le nouveau modèle de gouvernance sera remis en question par une croissance urbaine incontrôlée ;
- Quatrième implication, la dépendance des zones urbaines et littorales par rapport aux *lignes de communication* maritimes va se retrouver renforcée
- et, *last but not least*, cette urbanisation croissante va conduire l'OTAN à s'engager en zones urbaines.

Le FFAO, deuxième brique de ce processus militaire de transformation, a été lui publié au premier semestre 2018. C'est l'expression de l'avis des deux grands commandeurs militaires aux nations, et il va permettre de déterminer les scénarios probables d'engagement ainsi que les implications

militaires. Dans la suite, de la cinquième implication qui a été dégagée dans le SFA, le constat du FFAO est très clair, je le cite : « *un pourcentage croissant de conflits armés se déroulera probablement dans un environnement urbain, cette tendance existe déjà et le développement de l'urbanisation va seulement exacerber la probabilité de l'intervention de l'OTAN dans les opérations urbaines. Ainsi la ville, avec ses infrastructures et ses systèmes, va devenir la cible de l'action ennemie*. »

Le FFAO va dégager deux axes de réflexion sur le scénario probable de demain : le premier c'est qu'il va falloir renforcer la protection de la ville, il va falloir la durcir, et développer sa résilience ainsi que le maintien de ses capacités techniques pour faire fonctionner ses systèmes, ses systèmes d'énergie, ses systèmes d'approvisionnement en eau, en aide humanitaire, en santé, etc. Deuxième axe de réflexion, qui sera retenu, c'est que la numérisation de l'espace de bataille et le développement des systèmes autonomes pris sous l'angle de la dissociation entre le déploiement des systèmes d'arme et la présence d'un opérateur sur le terrain, donc ce développement des systèmes autonomes va permettre à l'adversaire de s'engager et de contrôler des espaces urbains de plus en plus importants, avec des forces de plus en plus réduites.

À partir de ce constat, quelles sont les capacités essentielles à développer pour opérer dans un environnement urbain ? Des travaux de réflexion ont été menés autour de huit grandes fonctions interarmées et il est apparu très rapidement, que les nouvelles technologies et en particulier l'intelligence artificielle (IA), vont être amenées à jouer un rôle central. Parmi les évolutions les plus significatives, je vais en citer quelques-unes :

Si on prend la fonction C3 : « *Consult, Command and Control* », les structures de commandement sont appelées à être beaucoup moins pyramidales, avec des capacités à louer des ressources très rapidement au plus bas niveau et une capacité à agréger et à désagréger des forces. On voit bien que dans ce nouveau modèle l'AI (*artificial intelligence*) est amenée à jouer un rôle central, en particulier grâce à la vitesse d'analyse et de décision qu'elle va pouvoir proposer.

Dans la fonction *intelligence*, renseignement en français, la ville va avoir une capacité unique à produire de la donnée, Michael Meier l'a évoqué hier. La somme d'information et des données disponibles, un instant donné, ne permettra pas quel que soit le nombre d'opérateurs humains disponibles, d'en tirer des renseignements actionnables à temps. La seule manière d'envisager une capacité de traiter cet afflux massif des données pour délivrer des renseignements actionnables, ce sera bien de recourir à l'intelligence artificielle.

Dans la fonction protection la force, « *force protection* », l'évolution qui sera probablement observée, consistera à avoir une capacité à mettre en œuvre des forces et des états-majors beaucoup plus dynamiques qu'aujourd'hui et beaucoup plus réduits en effectif. Afin de réduire leur empreinte et la capacité de l'ennemi à les localiser, la concentration des forces devenant un élément de vulnérabilité critique face aux capacités de ciblage de l'ennemi.

Enfin, les fonctions d'information opérationnelles et d'opérations civiles ou militaires connaîtront également une transformation assez importante avec une prépondérance de la sphère numérique et un nécessaire capacité à comprendre très rapidement l'ensemble des informations disponibles et d'y répondre également à temps.

Voilà alors, en conclusion, à partir du constat que les villes vont absorber et disperser très rapidement les forces - il apparaît que l'OTAN ne sera probablement pas à la mesure de mettre sur pied et de déployer une force massive pour le contrôle des villes, - la tendance sera à l'empreinte militaire la plus légère possible dans l'environnement urbain, empreinte militaire dont on a vu les possibles caractéristiques rapidement à travers une revue des fonctions interarmées.

Je vous propose maintenant d'aborder dans un second temps les questions juridiques qui risquent de se poser dans 15 ans, en 2035. Pour moi il y a deux grandes familles de questions : d'abord les questions qui émergent aujourd'hui et dont on sent qu'elles ne sont pas encore mûres, il s'agit des questions de responsabilité en coalition et des questions de combinaison entre les *policies* et les concepts qui offrent des protections de plus en plus larges, par rapport au socle, qui lui est stable, des obligations juridiques des Etats. Et en suite, dans un deuxième temps, il y a les questions qui relèvent plus largement la prospective, en particulier les questions qui concernent la coexistence de cadres juridiques différents et complexes lors d'une même action, et les effets de nouvelles technologies.

S'agissant de la responsabilité en coalition. Aujourd'hui l'Etat est responsable des violations du droit international humanitaire, le DIH, par un partenaire non-Etatique, si l'Etat a le contrôle effectif sur les opérations militaires, ou paramilitaires, au cours de laquelle les violations sont survenues, c'est ce qui ressort du fameux arrêt Nicaragua vs USA (CIJ, 27 juin 1986).

Demain, probablement d'une manière plus marquée qu'aujourd'hui, les opérations vont se dérouler en coalition et disposer d'un processus efficace d'établissement de la responsabilité au sein de la coalition sera un facteur clé de succès. Il s'agit d'un sujet extrêmement sensible, qui touche à la

souveraineté des Etats, et au centre de gravité de l'OTAN, c'est-à-dire, sa cohésion politique.

Alors, non que ses dispositifs actuels ne donnent pas entière satisfaction, mais plutôt la spécificité de l'environnement urbain va complexifier radicalement le tableau. La densité de la population, la possibilité de manipulations délibérées par l'ennemi ou par ses proxys rendent encore plus probable les violations du DIH, et complexe l'identification de ses auteurs.

Ainsi, au sein de l'OTAN, les Etats aujourd'hui ont une conscience de plus en plus claire que les commandants des forces opérationnelles, les « *joint force commanders* » et les Etats qui déploient les troupes, les « *troops contributing Nations* », pourraient être associés voire tenus responsables des actions d'autres acteurs, quand bien même ils seraient hors de leur contrôle. C'est bien pour adresser ce constat qu'il importe de perfectionner le processus actuel, afin de faire face à la complexité des engagements en zones urbaines.

Second sujet, actuel, mais qui n'est pas encore complètement mûr, les relations entre les obligations juridiques des Etats membres et les *policies* et les concepts. Dans l'OTAN il existe des *policies* et des concepts tels que la *protection of civilians* « POC », ou « human *security* », *policies* et concepts qui proposent une protection plus large en s'appuyant sur une approche souvent beaucoup plus globale. Eh bien, selon moi cette relation doit être finement analysée et mesurée (afin d'être expliquée et coordonée) pour dissiper les possibles confusions au niveau des commandements opérationnels. Cela concerne les questions évidemment cinétiques tels que le « *reverberating effect* » et là j'ai bien entendu l'argument historique du Professor Gaggioli, ou l'emploi d'armes explosives en zones urbaines, mais aussi d'autres sujets non cinétiques, tels que les réfugiés, les demandeurs d'asile, la protection des biens culturels, la protection d'environnement, et là on rejoint bien les préoccupations sur le long terme, ou encore, l'indemnisation des dommages, qui est un sujet qui est loin d'être aussi simple qu'il y paraît en coalition.

Au final, il va s'agir d'assurer la bonne intégration de ces différents éléments dans le processus de planification opérationnelle, sans aller au-delà des obligations des Etats, parce que, pour moi, c'est aussi au niveau des commandements opérationnels, qui sont la rotule des transmissions entre les intentions politiques et stratégiques et la mise en œuvre tactiques, que se joue une partie de la protection.

Enfin, *last but not least*, je voudrais aborder maintenant les questions juridiques prospectives, j'en ai retenu deux. La première concerne les

formes possibles des engagements en zones urbaines de demain. Une des caractéristiques claires de ces engagements sera certainement qu'il y aura une combinaison de différentes opérations simultanément : des opérations humanitaires, tels que la prise en compte de flots de réfugiés, des opérations de « *law enforcement* » (maintien d'ordre), des opérations cinétiques de haute intensité. Toutes ces opérations se dérouleront de manière adjacente et ça va nécessiter des forces extrêmement agiles, réactives, des forces déployées sur le terrain jusqu'aux états-majors. Je pense qu'il s'agira *in fine* d'être capable, le plus rapidement possible, de qualifier les différentes situations d'un point de vue juridique. Comme l'a souligné le Professor Venturini, cette tendance lourde aura probablement un effet important sur le processus d'éducation et d'entraînement. Il s'agira de toucher une audience beaucoup plus large et également de diffuser des connaissances beaucoup plus entendues. Le temps de séances d'informations sur les règles d'engagement avant le déploiement me semble révolu. C'est un peu la contrepartie de la diminution des volumes des forces engagées.

La seconde question juridique concerne les conséquences de l'irruption de nouvelles technologies dans les grandes fonctions interarmées. Cette problématique, que nous entrevoyons, a déjà été largement évoquée hier, lors des panels sur les systèmes d'armes autonomes et celui consacré à l'intelligence artificielle. En résumé, il s'agit de l'emploi d'un système autonome piloté par l'AI, elle-même alimentée par des *big* datas, avec un système d'apprentissage déterministe ou non déterministe, je vous renvoie à l'exposé du Professeur Chatila.

Ces systèmes autonomes vont soulever des questions d'opérabilité juridique, que je n'aborderai pas ici. En revanche, la problématique de la responsabilité est nouvelle et centrale.

Au niveau de l'OTAN, peut-être plus que les systèmes létaux autonomes dont la CCW, au niveau des Nations Unies, s'occupent parfaitement, les préoccupations aujourd'hui vont concerner les systèmes de commandement, les systèmes de C3 qui seront massivement impactés par l'arrivée de l'intelligence artificielle. Si on se réfère rapidement à la boucle OODA (*Observe, Orient, Decide, Act*), qui modélise la prise de la décision humaine, on s'aperçoit globalement que la seule fonction *« Decide »* reste encore aujourd'hui à peu près l'apanage exclusif des décideurs militaires, et quand on parle de responsabilité il s'agira bien de déterminer *a posteriori* comment la décision à l'origine d'une infraction au DIH a été prise. Souvent c'est à la suite d'une computation des données par des systèmes de systèmes entièrement ou partiellement autonomes. Ces

notions de prévisibilité et « *d'explicabilité* » des décisions vont devenir centrales si on veut éviter une dilution du lien causalité, qui est essentiel dans l'établissement de la responsabilité du décideur humain. Or c'est bien l'établissement de ce lien de responsabilité qui est une garantie essentielle de la mise en œuvre de la protection.
J'en ai terminé.

New technology and the preparation of urban warfare: what prospects for active and passive precautions?

Susan ESCALLIER
Brigadier General, Head, US Army Legal Services Agency

It is my great honor and pleasure to be here today. I would echo at the outset what my partners have also said, that I am here to share my individual thoughts. The policy of the Department of Defense is reflected in the Department of Defense Directives in a DoD (Department of Defense) Law of War Manual. It is especially an honor to participate in the Round Table as we recognize the 70th anniversary of the Geneva Conventions. I am Brigadier Susan Escallier from the US Army Judge Advocate General's Corps. This is my first time in Sanremo and I now fully appreciate the "Spirit of Sanremo". That spirit, to me, is one of hope and optimism, of open dialogue. The U.S. Army is so committed to the important work done here in Sanremo, that we send one of our top NSL legal advisors here for a year-long fellowship to learn from our partners and the world. That program is now in its sixth year. I could not be more proud of our Fellows or the support IIHL has given to them.

The organizers asked that I discuss "new technology and the preparation of urban warfare. What prospects for active and passive precautions?" I view our approach to using new technologies in urban warfare in the spirit of Sanremo, with hope and optimism but also, as the drafters of the Geneva Convention no doubt were - being informed by pragmatic considerations and mindful of the intersection of the law and realistic assessments of current and future conflict.

I want to start with the most fundamental of statements - the basics matter. In this case I mean the Law of Armed Conflict, or International Humanitarian Law (IHL), a term that is also often used. As such, when we must fight, the principles and rules of LOAC are our guide. In accordance with the principles of necessity, distinction, humanity,and proportionality, we all aspire to eliminating civilian casualties and eliminating the destruction of civilian objects, but that is not the law or always feasible in practice. Again, this should not surprise you coming from an American Soldier. And, I don't mean to come across as cavalier, but if yours is the profession of arms, and if called upon, the goal is to win and that often

means, even with our best efforts to minimize civilian casualties; they will still occur.

It was not long ago when the world was watching with great discomfort North and South Korea come close to armed conflict. Commentators, policy makers and strategists alike provided sobering predictions of what a war between the Koreas would do to the civilian population centers. That served as a wake-up call for many to think through the next fight and the human toll. LOAC and IHL are sufficient for the full range of conflict - the conventions that we honor with this 70th anniversary celebration were written with full understanding of the devastation of urban areas such as Dresden, Tokyo and London. And we are here as practitioners to ensure that the principles apply to any future conflicts.

As our National Defense Strategy identifies, and US Army senior leadership has stated, we are shifting focus away from two decades of counter-terrorism and preparing for the next fight - a peer-to-peer, or near-peer conflict. When considering how to man, train and equip for future conflict, we must be mindful that we have operated with policy layered on top of law - between ROE and LOAC. For example, legal advisors individually advising every target engagement authority has been a reassuring check for commanders in recent engagements, but in a fight for survival across multiple domains there will not be time nor manpower sufficient to expect a legal review for every use of force - policy required that process, not the law. (As an aside, there is an excellent article written by then DoD General Counsel Jennifer O'Connor who described the operation of a targeting cell during her visit to Baghdad). Instead, service members and commanders (who will be dispersed), to be ready for the next fight, must understand what LOAC requires and how to adhere to LOAC immediately and likely without the assistance of an on-hand LEGAD. Similarly, pages upon pages of ROE - something much of the forces fighting today are accustomed to - won't exist. In a war for survival, service members must understand and know how to immediately apply the basics, the LOAC.

We train our Soldiers in LOAC and these principles are incorporated and evaluated at our combat training centers. We frequently partner with our allies as we go through these crucible exercises and LOAC is factored into the scenarios.

At the outset I mentioned I view new technologies with hope and optimism. In short, we, along with our Allies and partners, are developing technologies to help better achieve the aims of LOAC, to increase our lethality and survivability to be sure, but always in compliance with LOAC.

What potential do the new technologies possess for our future? The ability to gather and synthesize data faster; better situational awareness - the time and space to understand the situation before taking action; and, the ability to provide our forces flexibility. For example, if an autonomous vehicle can replace an infantry squad in scouting a route through enemy territory, the commander can preserve his force and penetrate deeper. In addition to collecting valuable information for the commander to make a decision, if the autonomous vehicle does draw fire, the exchange produces at least two additional benefits: 1) potentially eliminating the need to return fire (because a human is not in danger), and 2) helping the commander identify the real threat without directly exposing humans. Or, in LOAC parlance, the use of new technologies will help a commander exercise additional feasible precautions to distinguish friend from foe.

What are these technologies? They are tools for the commander, tools for our forces to better engage the enemy more discriminately, which help them end fights sooner. States (in accordance with *jus ad bellum*) decided these wars were necessary. If we must fight, shorter is better—and doing so with greater precision is better. The new technologies we are working on are designed to accomplish those objectives. Now, as I mentioned, the US is preparing for a peer or near peer conflict which means the US may not have the technological advantage against an adversary. In fact, our Multi-Domain Operations Concept anticipates that every domain (land, sea, air, space, cyberspace) may be contested. Lack of overmatch does not change our obligation to comply with LOAC, it merely changes what is feasible. If our technology is rendered ineffective, we will still need to fight and to apply LOAC with reasonable available information.

And as we imagine the future world, we must be mindful that the world itself will be full of artificial intelligence and autonomous means to accomplish many tasks. It will be an autonomous battlefield with potential for logistics, maintenance and other functions to happen enabled by artificial intelligence and the world itself may see similar changes.

Because we do not know what the future will bring, the US trains and exercises across the full spectrum of contingencies. We run massive training exercises around the globe testing our ability to operate in the dark (analog) - this includes how we will operate under ground in tunnels and caves; we have bodies such as the Defense Innovation Board, which brings together industry experts, government officials and legal advisors as well as policy makers, who look at new technologies for functional and legal feasibility. Recently, the Defense Innovation Board ran a mock weapons review (Article 36 Review) assessing systems with various levels of

autonomy to understand the appropriate level of human judgment necessary for a given platform. They worked through these issues to identify what technology can do to provide increased lethality to our forces and also better enable a commander to comply with LOAC. All of this directly impacts the conduct of hostilities in urban areas.

This level of attention is not new to technologies like AI and autonomy, but we are cognizant of the perceptions and concerns surrounding AI and autonomy (some not based at all on fact or law). The current Chief of Staff of the Army, General McConville's top priority is people. Ensuring our forces are equipped with the tools they need to prevail lawfully (readiness) and adhering to LOAC to protect people are all part of this discussion. So too is ensuring that the public is informed with facts. Judge Advocates factor heavily into these tasks and we are remaining ready for whatever the future holds.

It bears repeating. States, and their armed forces, are comprised of people. People are running these exercises and interrogating these systems. People will employ this technology. The armed forces of states developing these technologies to win wars are also taking great pains to protect people.

Hopefully, the days of indiscriminate attacks on cities are a thing of the past, especially given the tools available to the many responsible armed forces around the globe. The bar is rising for what is a "feasible precaution" under the circumstances. Commanders have more information available to them to make better decisions. With that said, however, if the lights go out and the technology fails, the fight does not stop. Moreover, the analysis for what is a "feasible precaution" will change in the information denied environments we will certainly face in a peer to peer or near peer conflict.

LOAC exists to provide hard limits, to be sure, but it also allows flexibility for States, through commanders of their armed forces, to do what must be done - violently at times - under the circumstances. The technology in development, and the technology not yet even imagined, must facilitate the quick and humane cessation of hostilities. I am optimistic we won't see the "terminator" or its offspring any time soon. But technological advances will change the battlefield, and this isn't a two-hour Hollywood film. What we will see are commanders with tools enabling them to make better and faster decisions with more and better data. Maybe that looks like a microscopic electronic device looking at people in a building to distinguish civilian from combatant. Maybe it's an algorithm that can measure electricity consumption or track financial transactions unique to enemy forces. Maybe it is a longer lasting battery that helps a soldier on the ground maintain better communications with her commander enabling her

to obtain as much information as possible from her higher HQ before executing an assault. We can all think back into history to battles that did not need to be fought - but for a delay in a communication. September 1, 1939, marked the first day of WWII. What technology would have made that war less bloody?

We are better equipped today to handle those issues and we are far more deliberate in our analysis of technology employed on the battlefield than ever before. Regardless of the technology, we must focus first on the basics, with LOAC. As I mentioned above, people will drive the application of technology and determine how or if it may be employed on the battlefield be that urban spaces or open oceans. Just as the drafters of the GCs were reflecting upon and working to address the atrocities they witnessed , people today must anticipate the impacts of technology now and in the future and be mindful of LOAC and ever aware of the changing environments where conflict occurs.

Risks in using new technology in urban warfare – and additional steps States should take to avoid civilian casualties

John AMBLE
Editorial Director; Co-director of the Urban Warfare Project,
Modern War Institute at West Point

Before I begin, I would very sincerely like to thank the organizers of this event for the invitation to be here and to share my comments. I don't work for a humanitarian organization, like many or most of you do. And I'm not a lawyer, like many or most of you are. But as a result of that, I have learned an extraordinary amount during the panels these past couple of days and have really enjoyed the conversations I have had the chance to have with several of the people convened here on the sidelines. So, I hope I can repay that and that I can offer some value to the roundtable's proceedings with my comments.

Since I'm employed by the US Army, I'm obligated to note up front that, firstly, while I'll make some comments that I think are generalizable to many military forces, my perspective is based very much on the experiences of the US Army. As a result, for the most part, my remarks really focus on ground combat forces in urban environments, and don't really engage very much with the question of airpower and cities, which is an important one and an interesting one, but is beyond my scope today. And secondly, while I shall talk about the US Army and the US military more broadly, my comments are not official positions of the Army, the Defense Department, or the US Government.

I believe I was invited here to talk about the state of thinking about military operations in cities within the US military and more generally to offer a military operational perspective.

I work for a US Army organization, based at the United States Military Academy at West Point, called the Modern War Institute. I am also here in my capacity as co-director of that institute's Urban Warfare Project. My organization is actually quite small, certainly by US military standards. We're a team of just a few people, but we have become, I think, one of the foremost and at least, perhaps, the most prolific outlets within the Army in terms of thinking about conflict in cities. That is, in part, because it's something we see as really, really important. But equally, I think it also

says something about what is, to be frank, the pretty limited institutional attention paid to cities, not just in the Army or the US military, but I think in many modern military forces around the world.

There has certainly been a notable increase in emphasis in the US military on urban warfare and the idea of military operations in cities - at least in terms of senior leader statements. However, if a US Army unit were called upon tomorrow to fight in a dense urban space, there would be a range of challenges unique to cities that that unit would be woefully unprepared for.

That has important implications, which I want to discuss and which we'll get to, but first, it's important to ask: Why? What explains this lack of preparedness to operate in cities, and more broadly, this lack of attention to them? And when I say "lack," I don't mean entirely lacking. I mean less attention than you might expect given global demography, future trends, and how badly urban conflict often turns out.

So, why? One of the most important points to keep in mind in answer to that question is this: For most of history, militaries have fought FOR cities, but not IN them. As a result, state-based militaries have been organized, trained, and equipped to operate in large, open spaces with plenty of room to maneuver.

We can think, for example, of the Fulda Gap - has anybody here been to that part of central Germany? If so, you will have seen the perfect example of what I mean - wide open plains devoid of complex terrain and largely devoid of non-combatants. The Fulda Gap was, of course, the planning scenario, in terms of conventional operations, that governed the way that NATO militaries were organized, trained and equipped during the Cold War. But that wasn't some new function of a set of characteristics *unique* to the Cold War. It has been the way militaries have been organized, trained and equipped for a long time, certainly as far back as when maneuver warfare came into its own during the Napoleonic Wars.

That trend extends even further back, at least to when the Romans professionalized military service and discovered that mass was a remarkably influential determinant of battlefield success. Soldiers, throughout much of history, might have lived in villages, or towns, or cities, but they would be formed up and would march out to battlefields that were not in cities, where, it's also important to note, they would fight with increasingly heavy weapons and armor.

So that historical context is really important to remember. The fact, then, that today, military forces are not organized, trained or equipped for cities should really, maybe, not come as such a surprise when it's looked at

in that broad, historical context. And yet, despite all of this, military forces still have repeatedly found themselves operating in cities including Aachen, Stalingrad, Hue, Fallujah, Ramadi, Mosul and Marawi.

This is, as I'm sure you all know, not an exhaustive list of urban battles. But what do these examples from the past three quarters of a century or so have in common?

Firstly, they involved state-based ground combat forces that were NOT optimized for cities. Secondly, they were incredibly destructive. Thirdly, one of two conditions existed either:

a) They were fought by two sides, neither of whom really wanted to be fighting in that city—Stalingrad is a very good example of this, of how a city can become a terrifyingly destructive battlefield almost, not entirely, by accident. That's the first condition.
b) Or one side was demonstrably weaker and chose the city as a battlefield because of its leveling qualities, stripping the stronger party of many of its advantages, in terms of armaments, technology, and more.

So, by looking at history, we can begin to understand why modern military forces are not, again, as I said, organized, trained or equipped to operate in cities.

But also, we've seen that despite this, urban centers have repeatedly pulled military forces in. And those cases, when that happens, as I said, had a few common characteristics.

Firstly, I want to focus in on the second characteristic - that they were incredibly destructive. What explains this? Why is warfare in cities so destructive? Answering that question is clearly a necessary first step toward diminishing that destruction. That is, of course, an important goal of many of the organizations represented in this room and I think of most military forces, as well.

We heard in a comment yesterday during one of the panels from a gentleman who said that in teaching here in Sanremo, one of the key points they try to get across is that by complying with IHL, a military commander will be more militarily successful. I certainly agree with that, and I think that becomes most apparent when you focus on longer-term, strategic military objectives. We all have an interest in limiting the effects of military operations on civilians, physical property, infrastructure, cultural property - all the sorts of things that you find a lot of in cities.

So, back to the question: Why are military operations in cities so destructive?

Remember, I've used this phrase several times now: organized, trained, and equipped. So, if we think in those terms, organizational changes are the most difficult to conceptualize in terms of how they might make military operations in cities more effective and less destructive. But really, the impact of the way militaries are organized is probably the least of the three. Whether a force has squads with 9 soldiers, or 10 soldiers, or 13 soldiers - whether an army is built around the division level or the brigade level - these are important organizational questions for the military, but probably not all that impactful on how operations are conducted in cities.

Secondly, I would like to focus on training. Urban terrain requires unique types of training. What is required to do the basic soldier tasks of "shoot, move, and communicate" in a dense city is very different than what is required in comparatively open terrain. To do those basic things with so many civilians in the battlespace requires specialized training. Now, we've seen training improvements in recent years. Better methods, better facilities that more accurately replicate the complexity of a city, but we still have a long way to go.

The most common form of training for cities which we used to call MOUT - military operations in urban terrain - is really limited to a single task: enter and clear a room. It's a pretty straightforward thing, but when you actually break down the mechanics of a team stacking on a door, breaching, passing through what we call the fatal funnel, and each team member scanning his or her assigned portion of the room, while moving, while trying not to trip on the power cord or weapon or clothes or toys or whatever else might be in there, and making split-second decisions about the people in there, it's tough. Now doing room after room after room after room, which is how many urban battles end up, that's a persistent exposure to incredible complexity that, I think, is extraordinarily difficult to train for.

And that's just the tactical complexity - it doesn't even touch on trying to train soldiers for the strategic complexity of a city. A colleague of mine who has spent years studying urban conflict has a very good analogy. He says military operations in open environments is like playing billiards. You strike one ball and you pretty much can predict the other impacts, the follow-on effects. Conducting military operations in cities is like playing with 100 other balls on the table. It's virtually impossible to predict the second and third order effects and even beyond each tactical decision made, which can in many cases have strategic consequences.

So, that's organization and training, which brings us to equipment. These are the tools that a force brings to the fight, and to be clear, the issue of equipment is the most explanatory factor, I think, in terms of why

military forces struggle to be effective in cities and why the effects are so destructive.

The equipment that ground combat forces use - from small arms to artillery and from personnel carriers to tanks - are designed to maximize effectiveness in open spaces. Their limitations are on full display when they are used in the condensed spaces characteristic of cities. However, military forces must operate with the tools that they are given, and as long as those tools are not designed specifically with the density of cities in mind, those forces will struggle to maximize their effectiveness and their operations will be more destructive than anybody wishes.

Make no mistake, new technology will help. Greater precision, more adjustable payloads that allow munitions to be delivered with the minimum necessary firepower to meet the military objective. But it is difficult for me to conceive of any technological development or combination of technological developments, and I think we would make a mistake if we have that expectation about the future, that technology is going to solve the problems associated with protecting civilians on the battlefield.

I know our moderator on this panel has said he has non-lethal tools to keep each of us from speaking for longer than we promised, and I hope I'm not at risk of being the target of his bell, but I find it fascinating that he said that, because that means he is better equipped with non-lethal tools than the typical soldier deployed in an urban environment.

There are a lot of reasons for that. In part, it's a function of deliberate decision-making - remember, most forces are organized, trained, and equipped to operate in open terrain without civilians on the battlefield. There's not much need for non-lethal tools in those areas.

But there are other reasons, as well. I spent 15 months deployed in east Baghdad in 2007 and 2008. Our area of operations spanned several sectarian fault lines that crisscrossed the city. There were also a number of Shia holy sites in the region and every few months there would be a pilgrimage with tens of thousands of people marching through the city. Unfortunately, there were also places these marches passed that left these crowds vulnerable and they were, on a number of occasions, targeted by suicide bombers, car bombs and rocket attacks. Coalition forces had few, if any, non-lethal tools with which to disperse crowds and keep them from these sites where they were most vulnerable.

What would police forces in many countries, including my own, use to disperse a crowd? Tear gas or, in other words, CS gas. It is remarkably unpleasant. Has anyone in here experienced it? I have. But it generally works, as well. It's also banned for military use by the chemical weapons

convention. That convention is inarguably an achievement. But I still remember soldiers in Baghdad asking why we don't just use tear gas to disperse the crowds instead of seeing them targeted and seeing innocent civilians killed. For those soldiers, that's their only touchpoint with the chemical weapons convention and is a perspective that I think is important to keep in mind when we talk about non-lethal weapons and their utility on the modern battlefield.

I think I'm going to leave it there, although there's certainly scope to flesh those ideas out much more. The main point I'd like to leave you with is that, in terms of limiting the destructive nature of urban conflict, technology is going to be a help, and we as a military community and the humanitarian and legal communities are all going to have a role in making sure technology is leveraged to that end. But don't expect technology to solve the problems entirely.

Enhancing military effectiveness in cities and limiting the destruction caused by military operations is ultimately, I think, going to require changes to the way we organize, train and equip our forces.

VII. The prospects and pitfalls of digital technology in designing and delivering effective humanitarian responses

The impact of new technology on the ability of organizations to provide humanitarian assistance

Hovig ETYEMEZIAN
Head, UNHCR Innovation, United Nations Commissioner for Refugees (UNHCR)

I would like to start by saying that in the past two missions I had in Iraq - you have heard what happened in Mosul - I saw a good example of how we could use technology to change the way we did business there a little bit in terms of accompanying the return of Iraqis who had fled Mosul: starting a cash for shelter program, which used mobile money. That ability to use mobile money gave so much more flexibility to the people we are paid to serve. For them it was important to be able to take ownership of the way they did the return and the way they rebuilt their houses. That was a positive example of how we could use technology to help people have more ownership and more control over their lives.

Taking a step back in Jordan, by the time I left, we were using Iris technology for refugees to get their cash assistance by going to ATMs and scanning their iris and getting cash. Now, it's my first time living in a place like Switzerland. Here, I don't use my iris to go to the ATM and cash money. So, it was quite a step forward. These were positive examples. They are not without risks, but today I actually want to take a step back and focus on what is right before technology. I'm very happy that there's a very good presence of ICRC there, because ICRC has been quite the leader in dealing with this kind of responsibility, the responsible use of technology and how we can use technology, but also the pitfalls that it might represent and how we should be careful with it.

I wanted to take a step back and say that today we are close to 70 million forcibly displaced people between those who are internally displaced and those who are refugees. The number of refugees in the world is 438 times the population of Sanremo and 312 times the capacity of San Siro football stadium. It is quite a sizable proportion of the world and only 16% of these refugees are in developed regions. Most refugees and displaced people are actually displaced in neighboring countries, and actually suffer from major challenges.

For us, the main question is: “how can we provide solutions that cater to the wellbeing and dignity of forcibly displaced?” This is what we’re trying to solve when we talk about delivering services. How can our team, the innovation team, provide solutions? That’s the starting point. If I were going to leave you with one message today, I would want you to remember what we do in innovation. For us, innovation is not equal to technology, innovation is not simply technology. It is not only for the younger generations and it is not something only a few people can do. There are people who can confuse the terms innovation with technology. People build apps and use block chains and drones to deliver services and they think of that as innovation. Additionally, there are some other people who think that innovation is all about ideas, when, in fact, it is more about delivery, people and process. For us, innovation is about action. Without implementation, ideas are just ideas and ideas alone are not innovation. Responsible actions refer to respect of humanitarian principles and the ethical values of our organization. I think our colleagues from ICRC will talk about it, but I think the underlying theme for us is the “do no harm principle” and I’ll come to that later.

We can see that there are maybe four dimensions to innovation. One is product innovation that is, producing new things like products, for example, drinking water filters or a website, or an app. There is a process innovation, which changes the way in which we create and deliver, for example, Ford’s production line or the cash assistance programs. These are process-oriented innovations. There is also position and policy innovation, for example, recently Uganda allowed refugees to have access to SIM cards, so they made a decree and they changed the law. Now, refugees can have access to SIM cards using the refugee status condition they have. That was a change in policy which we see as innovative. And also, there is paradigm innovation and those are the mental models that we change, where we shift, for example, low cost airlines, that’s a paradigm change. We are changing the whole spectrum of what we do a little bit.

Our proposal for you today is to focus on innovation and not technology to deliver humanitarian response. Humanitarian needs are only going to grow and the resources available to us are not likely to match that need. We work in places where there is no running water, no electricity. Technology, therefore, cannot solve all the needs of all the people we serve. Technology can definitely not solve all the needs and has many pitfalls. But when solutions are centered on people, they become sustainable. This is the main premise of innovation. When the values that underpin our attitudes and behaviors as humanitarians drive innovation, we can better focus on our

efforts, and create more impact with less. And we're going to have to, because we don't have enough resources. Innovation is important for us because it makes us more agile, more open to collaboration and more effective for the people we serve. This will help us deal with future forced displacement.

So, there are certain factors and variables that govern displacement in the future. Climate Change is one of them as it exasperates displacement. It's one of the stress factors for countries with a history of protracted conflict. Somalia is a good example. In Somalia we are using artificial intelligence with piloting uses of artificial intelligence to predict movements of populations. This is one example of how we could be using technology positively, but we didn't jump to that, we wanted to predict displacements and, therefore, we identify the problem and we're only using technology as a means to an end and not an end by itself.

Another phenomenon is the mixed migration flows we have now. Look at what happens in Libya and the mixed migration flows to Europe, or what's happening in Venezuela. We increasingly have challenges related to mixed migration. The third element is the rise of extreme nationalism and the usage of, let's say, social media and technology to exasperate negative emotions and negative feelings towards vulnerable populations. Another element is increased surveillance by state and non-state actors, who use more and more personal data of individuals and there is a risk of using personal data of individuals who are extremely vulnerable.

So, for this reason, we want to share certain principles or ideas; prospects we consider are needed to design better solutions; and deliver appropriate humanitarian response. These prospects create value to those who are forcibly displaced and provide sustainability to the solutions designed and implemented. Those solutions could either use technology or not. So, what are those five prospects?

The first one is: question your assumptions and co-design with the people you serve. One of the common issues when designing any solution for humanitarian response, technology-based or not, is that people designing those solutions need to be humble enough to question their assumptions. When delivering a response, the first thing we need to do is to ask the people we serve if our solution is appropriate to their needs. And, ideally, and I emphasize this point, we should co-design any solution with them. Make them part of your planning and inputs when you're designing. I'll give you a negative example that I faced in Zaatari, in Jordan. We had a startup that came and wanted to use 3D printing for prosthetics; so, they said "we're going to bring our 3D printers and print prosthetics for disabled

people in the camp." My question was, "did you actually speak to the refugees in the camp? Did you speak to the health authorities in Jordan and establish?" There are specific protocols that govern the whole process of users of prosthetics. None of that was done. They thought they had found the solution, which was 3D printing and, of course, maybe 3D printing was great, but that is only a part of the solution. So that didn't really go far.

Maybe a good example was given by the UNHCR biometrics team regarding the use of technology. The team noticed that refugees were curious about what was collected from them, when we did our registration process with refugees. You have the same curiosity when you're in an airport, when the airport officials collect your fingerprints and you cannot see what is on the screen, as you're standing on one side of the screen and you don't know what's happening on the other side. So, when we were testing our iris scan technology, we realized that there was something missing, because the client was not really seeing the process. This is a minor tweak, but today our colleagues in the field have the guidance that when we take the iris scan of refugees, refugees need to stand side by side with our colleagues and, therefore, they can see exactly what's happening and we explain and we should explain to refugees the process of us capturing the data including the iris scan.

The second prospect is to adapt response to the appropriate context. Technology solutions should always consider the constraints and limitations of the context as well as cultural appropriateness for solution. I'll give an example. There's a tendency to do hackathons. These hackathons are events where developers and other experts in technology gather and in 48 to 72 hours they design the solution. We're not really big fans of these hackathons, because it feels that they are not involving those of concern. Those people who we are paid to serve, need to be at the center of a solution. And a solution cannot be the best solution unless it's tested together with our clients. And so a lot of these hackathons happen in a silo, in a bubble, and that bubble doesn't include the clients. I'll give an example of a very, what you might consider, a low-tech solution, namely, the Boda Boda Talk Talk. We basically had information gaps in a large camp. We put the radio, basically speakers, behind the motorbike, the motorbike would go around the camp transmitting the same message in a large distance and would stop along the way because people would listen and then have questions. So, we used low-tech to handle misinformation and lack of information. This is just an example of solutions that do not necessarily amount to the high-tech solutions that people usually think about when we talk about solutions.

The third prospect is design solutions with risk and mitigation measures. Another common pitfall and particularly concerning technology solutions is that some of the solution designers do not focus on risk and potential mitigations when they are implementing them. They focus only on the novelty of the solution rather than the potential risks the solution can create or the dependencies that they might generate. For example, the Distributed Ledger Technology (DLT) or what is now commonly known as blockchain. This has been considered one of the most promising technologies of the century. It's essentially a shared database filled with entries that must be confirmed and encrypted. The name blockchain refers to the blocks that get added to the chain of transaction records. Each data entry is dependent on a logical relationship to all its predecessors. Blockchain technology has proven to be successful in the case of financial transactions, and to tackle corruption. There are people who now suggest the use of blockchains for holding personal data record. Again, blockchain needs to be understood and studied before you're able to say that blockchain is a solution. So, I'm just inviting you to think about it. When we think about blockchains, we should be thinking first about the problem that we're trying to address and also look at the users of the blockchain.

One opportunity here is cash assistance. A lot of studies prove that cash assistance is a positive methodology used to provide assistance. However, we should constantly take into consideration risk and mitigation measures on cash, starting from the introduction of additional liquidity and, therefore, inflation, the security measures for distributing cash and then also, behind this, the management of the identity of those who are receiving cash, how we mitigate the risk with mobile operators who have access to some of the components of the identity. These are important aspects to consider.

We should be focusing on agency and accountability. Agency refers to the capacity of individuals and collective groups to act independently and make their own free choices. Accountability is about taking account of, giving account to, and being held to account by the people we serve. We have to break with the idea that we are best suited to design technology solutions or solutions in general, more than the people we serve. So, again, we must consider the issue of including our clients in the process of finding solutions.

Five: stronger policies that protect people, particularly data protection. We tend towards data protection, which is part and essence of the way we manage the confidentiality of the data of the people we serve. We do have a data protection policy and this is something that we want to emphasize. So, the stronger your policies are on data protection and the more elaborate

your systems, the easier it will be to use technology while mitigating the risks.

Basically, for us, these are the five prospects that we would want you to consider when designing and implementing technology solutions and maintain a response. One: question and validate your assumptions; two: focus on agency and accountability; three: design with risk and mitigation measures; four: adapt response to appropriate contexts; five: strengthen policies, especially in this set up on data protection.

The humanitarian metadata problem: 'doing no harm' in the digital era

Alexandrine PIRLOT DE CORBION
Director of Strategy, Privacy International

First of all, I wanted to thank very much the International Institute of Humanitarian Law and the ICRC for inviting us to be part of this session this morning. While we may seem to be an unusual ally as privacy advocates to be working with the humanitarian sector, these opportunities are important for us to identify common areas of concern to develop joint efforts because ultimately, we have the same vision: safeguarding people's dignity and autonomy.[1]

For many years, Privacy International has been exploring the deployment of new technologies and the use of data intensive systems in development and humanitarian programs[2]. And we've been exploring what it means to do no harm in the digital age, in an increasingly challenging context for humanitarian organisations: the pressure to be accountable and transparent by funders, to be efficient, the sustained needs that emerge from conflicts that last longer and also with armed conflicts being fought differently, as we've been talking about over the last few days.

There is no question that an advancement in technology, communications and data processing in capabilities have changed the way humanitarian assistance is provided. But we also need to be considering some of the challenges that are emerging as part of that. There has been a digitization and "datafication" of our societies as mentioned by Helen Durham on the first day of our Round Table and the humanitarian sector has not escaped that trend, but the question is, at what cost? And, do we have the information available to do that cost benefit analysis both in the short term and in the long term?

Technology is not neutral. Data is not neutral. Innovation is not neutral. Because those deploying and designing these systems are not. Despite some efforts by some organizations, and those efforts are growing, overall, unlike many other sectors, the humanitarian sector has not been subjected to the same level of scrutiny when it comes to proving the impact of some

[1] https://privacyinternational.org/strategic-areas/safeguarding-peoples-dignity.
[2] https://privacyinternational.org/topics/development-and-humanitarian-sector.

of the innovative solutions they are adopting on individuals and ultimately undertaking benefits and harms assessment. One reason could be the lack of a single common accountability mechanism for the sector as a whole. It's really interesting that for a sector that historically has been at the forefront of risk management and identification, these approaches and strategies have not seemed to be replicated with quite the same rigor as in the past for non-digital humanitarian intervention into digital humanitarian actions.

Before I start, similarly to what Hovig said, I want to highlight that, while aspects of this presentation will be technical and full of the sexy buzz words emerging from a broader hype around innovation, we must remember that behind every example, every anecdote, every reflection or news report that I will mention there are people and communities, who are directly impacted by the issues I will discuss today. When talking about datasets and data, there is a tendency to forget that, and we must not forget it.

So, here is a quick outline of my presentation; first, starting with a short introduction to the study that was mentioned before by Nils; as well as introducing the concept of "do no harm"; and then I'll be reflecting on the study and generally the work that we've been doing with the humanitarian sectors to highlight what we know and which steps we should be considering moving forward.

In terms of the humanitarian principle of "do no harm", there are various interpretations of it and they've been expressed in different ways, but I've chosen the one here by the ICRC, which reads that "a humanitarian organization needs to ensure that its actions do not have an adverse impact on, or create new risks for individuals or population". Much of the literature and analysis of the use of innovation, tech and data in the humanitarian sector have been exploring some of these issues through the lens of these principles and new developments to ensure that humanitarian organizations comply with it.

So, in terms of the study, you may ask yourself: why metadata? This is the title behind my presentation, which includes the humanitarian metadata problem, "doing no harm in the digital age". There's a well-known phrase which you might be familiar with by General Hayden, former director of the United States National Security Agency and of the Central Intelligence Agency in May 2014, when he said: "we kill people based on metadata".

His comment, in spring 2014, was the first to articulate the power of metadata, so succinctly and famously. The reason why this statement became so recognized and shocked many wasn't so much that metadata could reveal information, and I'll explain what metadata is in a second, that

was already well known, but it was the extent to which that information was being trusted and used to make life and death decisions. In the years that followed, there was more information that emerged from the breadth and the scope of global intelligence agencies, mandate, function, operations and oversight or a lack thereof. Some agencies were listening in on calls, monitoring or intercepting online communications and tracking individuals through a variety of digital and non-digital programs. That was the starting point of this particular collaboration between Privacy International and the ICRC to explore humanitarian metadata[3]. That is to say that metadata generated through humanitarian actions.

Now that I've said the term several times, I wanted to take the opportunity to explain what that is. Metadata[4] is a set of data that describes and gives information about other data. So just to give you a very simple example, when you write an email, the metadata would be your name as the sender, it would be the information about the recipient or recipients that you're sending it to, it would be the subject, it would be the timestamp or the location, which mail account you use to send it. It could also include whether you had an attachment, a photo, a video or other type. Whilst we are driven by assessing the generation and processing of metadata, the computer scientists in the room will ask why? Data is data, why make the distinction? And so that was also integrated over time in our research that it was not just metadata, but it was also about content and it wasn't just that, but we need to start thinking about categorizations of data: declared, inferred and intent. They provided us with a framework to develop our problem statement, our thread of concerns and also to explore the implications of new technologies and data processing capabilities. While there are many issues to explore from biometric registration of refugees, to the centralization of personal data in the form of digital ID systems, to the use of drones, AI machine learning, our study focused on three areas. We looked at traditional forms of messaging and messaging apps, cash transfer programs and social media platforms.

In order to give you a quick example of what we were looking at particularly, I've chosen to give you an example of social media. For humanitarian organizations physical access to people affected by crisis isn't always possible and it's become increasingly difficult. We are seeing that

[3] https://privacyinternational.org/report/2509/humanitarian-metadata-problem-doing-no-harm-digital-era.

[4] https://privacyinternational.org/video/1621/video-what-metadata.

humanitarian organizations are using social media for a variety of reasons, whether to provide key information, to gather information or to support their programs. But if we think about how social media platforms operate, and we look at the different categories I've mentioned, metadata, declared, inferred and intent data, this is the sort of information that we were looking at. When it comes to social media, metadata would be your user profile ID, location, friends, likes, comments, content access, links clicked on. The declared data is the data that you would be providing as a user, where you know that you have given that information. It could be things that you provide when you sign up, your name, gender, date of birth, depending on the service. But there are also things like inferred data. So, based on your interactions online, the data you provide, or that others provide within your networks, other pieces of information could be inferred. For example a study by Cambridge University[5] in 2013 revealed that they were able to use an algorithm to infer information about users from their data and their online engagement including users' sexual orientation, satisfaction in life, intelligence levels, emotional stability, alcohol use, relationship status to name a few.

Then you would have intent data. So, this is data being used to predict behaviours, to identify trends. And there I want to point to a study which revealed that based on 10 Facebook likes or 300 likes, depending on the side of the study, they were able to make assumptions about users. For example, participants with high openness to experience tended to like things like Salvador Dali, meditation or Ted talks.

But here we need to also start thinking about the risks. And these are some of the things we highlighted through different scenarios in our study. What are the risks when you're able to, or at least try to, predict people's behaviour, their preferences, other personal details, also to create inaccurate profiles? A lot of this information is based on data but depending on the quality or other sources of data that are fed into it, it doesn't necessarily mean it's a true reflection of who you are as a user.

It can lead to surveillance, false identification and targeting. And then you might start asking yourself, why am I mentioning all these things? Why is that problematic? I want to give you a few examples as to why that's problematic, where the attack surface created by social media platform and the inherent vulnerabilities of the designs of these platforms

[5] www.cam.ac.uk/research/news/digital-records-could-expose-intimate-details-and-personality-traits-of-millions.

are used against us. To give you a few examples, we've had a few European governments as in Austria and Germany using social media accounts in decision making for asylum procedures. The power of metadata and other social media information is also reflected by policies like the one of the Trump administration asking people coming into the US for their social media accounts. But you can also see it in other sectors like the Fintech industry, where profiles are created to see whether you'd be a reliable person to give a loan to. So, these are some of the issues that we addressed in the report that we contextualize to the humanitarian sector.

Moving forward, for the next part of my presentation, I'm going to move slightly away from the study to broader issues that we have been looking at with the humanitarian sector. The first one, in terms of what we know, is that humanitarian organizations both drive and depend on the data generated and the process through their different activities. It could be by choice because of the way they design their programmes, and some examples were given in the presentations in terms of having a digital ID system, a cash transfer program requiring information from beneficiaries, but it is also because of the infrastructure the humanitarian sector relies on. It is often provided by third parties because they don't develop their own tools. Yet, what we've noticed is that there's very little regulation, awareness, or training and not enough is being done to identify and to mitigate the risks created when using systems and platforms designed and managed by third parties.

The other component of what we know and that's not just limited to the human sector, is that governments have vast, unrestrained and unaccountable powers that threaten freedoms and rights. They no longer concern themselves with individuals, but entire populations, groups or regions can be and are being placed under surveillance. The humanitarian sector has been a victim of some of those practices. In 2013, journalists reported that a list from whistle-blower Edward Snowden, included information that UNICEF, *Médecins du monde*, UNDP and other humanitarian organizations were the surveillance target[6] of British and American intelligence agencies.

The other component we can't ignore, and I was mentioning that humanitarian organizations don't operate in a silo and depend on third

[6] https://privacyinternational.org/news-analysis/1023/theres-no-good-reason-spy-agencies-snoop-humanitarian-groups.

parties, is that companies routinely exploit people's data[7] for their own advantage. They often do so with a lack of transparency and accountability because their aim is to generate profit and fuel their own business model based on data exploitation. Their business model is not to protect people.

A key component that's been focussed a lot in our discussion in the last few days has been around security, not only security of people or national security, but also the security of our infrastructure. IT systems are inherently vulnerable to intrusions or data breaches. There are numerous high-profile examples and I've put a few here, of breaches happening every day and there are millions more not being reported. And even the most well-resourced governments in the world are unable to protect their most sensitive data sources. For example, the US civil servant database of government employees was breached a few years ago. And then more recently, as you can see at the bottom, Facebook admitted[8] that the phone numbers of 145 million users were leaked. You'll never have 100% security and safety, but you can make the job of your adversaries much harder.

So, as I go into the conclusions, we need to start thinking about what needs to be done differently and not just the humanitarian organizations, but also the other actors within the ecosystem, the beneficiaries, their implementing partners, third parties, governments, but also funders. When you contextualize everything I've presented so far, how data and technology is used for the delivery of humanitarian aids, and we consider today's ever-growing digital world, this means that more and more people receiving humanitarian assistance have the potential of being exposed to unexpected threats. And unless measures are taken to address some of these issues, it will be increasingly challenging to sustain the impartiality, neutrality and independence of humanitarian action.

And so, I wanted to highlight four areas which need to be considered moving forward. The first one is to really take a step back and acknowledge the implications of the decisions that are being made, particularly at HQs, because there's a huge division between the decision making at headquarters and then the implementation in the field. And on that first point we need to think about taking a step back, questioning *why* are we using data and tech, versus *how* can we make use of data and tech? We can

[7] https://privacyinternational.org/blog/2536/companies-must-do-no-harm-humanitarian-sector.

[8] www.forbes.com/sites/daveywinder/2019/09/05/facebook-security-snafu-exposes-419-million-user-phone-numbers/.

make use of the advancement in technologies, but we need to understand the needs as well, and whether the solutions identified are really going to solve those problems. As part of that assessment, we also need to think about the legal regulatory and ethical obligations, that humanitarian organizations have. And I mentioned ethical obligations as well because many organizations, including the ones represented here today, have privileges and immunities, which means that they don't necessarily have any legal obligation to particular jurisdiction and it really comes down to their moral and ethical obligations to protect their beneficiaries. The other core element when it comes to risk, is implementing security and risk management, protocols and also incident reporting. In the study that we did on humanitarian metadata, not one single evidenced incident of reporting of a breach was given to us. It is incredible that in all of the organizations we've spoken to, no one had recorded officially a breach of a leak of personal data. Given that they happen every single day, even to the most well-resourced governments, it's a clear indication that we're not looking at some of these issues in depth.

The second one is around preparing to fail well. There is a need to invest time and resources in this process as humanitarian organizations prepare to deal with the good and the bad of using innovation. Does that also mean changing their structures? How they operate? The kind of skills they need internally and not just having the IT guy who sits at the back of the office? Humanitarian organizations don't operate in silos. So, there is a need for collective action. They need to improve, across the sector, data protection, privacy, insecurity, safeguards, and to make similar demands on those they work with, both internally and externally, to reduce the attack surface across the sector.

The last point I wanted to make, and it's more of an advocacy point[9] than an academic one, is that humanitarian organizations have a huge advantage in terms of how they're seeing some of the negative impacts of some of these innovative solutions being either presented or implemented by them. But there's a need to re-focus and think about having the individual at the centre and in control and with dignity. And that also comes with putting pressure on funders. That's something that keeps coming up when we speak to humanitarian organisations: "Our funder wants us to use biometric" or "Our funder wants us to use block chain", or

[9] https://privacyinternational.org/news-analysis/2535/do-no-harm-digital-age-privacy-and-security-cannot-be-ignored.

“Our funder wants us to use artificial intelligence”. Often, they are making those demands and conditions to receive funding, but without understanding the risks, nor resourcing the humanitarian organizations to mitigate them.

There’s a real opportunity here for humanitarian organizations to be heard and shape the direction, but also to think across the sector, to share their knowledge of expertise. We’re not a tech phobic organization, and our advocacy in the work that we do is, how can we use innovation in the most responsible way? How can we provide actors who are using these tools with information, knowledge, and expertise, but importantly the processes and protocols so that they can make informed decisions and learn to fail well?

Thank you.

The use of new technology in humanitarian action: a challenge for data protection and the principle of independence?

Martin STANLEY SEARLE
Associate Research Fellow, S. Rajaratnam School
of International Studies, Nanyang Technological University

Ladies and Gentlemen this past June the Sana'a-based Houthi authorities in Yemen were reported to have refused WFP access to areas under their control due to concerns over the "national security" implications of WFP's biometric registration process. While some dispute the rationale given by the Houthi administration for this decision, let's take it at face value for a moment and consider the critical question it raises. What should we do if embracing new technologies dramatically improves the quality of the aid we can give, but reduces the number of people we are able to give it to? By "we" I mean military representatives, lawyers and aid workers, because, as I shall argue, each is involved in some way in this issue.

I'm asking this question here not to answer it. It's not really my question to answer. I'm asking it because my presentation, and the paper it's based on, concerns why this question needs to be asked.

I've been periodically presenting and revising this paper for the best part of 18 months. I think the best thing I can do in the context of this session is to focus on two things. I'm going to first highlight some key points within the ongoing discussions of both digitized data and independence in humanitarian work. I'm then going to tie these together using something called the "capability-vulnerability paradox" – a concept I am borrowing from military scholarship on cyber-based weaponry.

Independence is one of four main humanitarian principles, the others being humanity, impartiality, and neutrality. These are not simply abstract values, or even purely ethical ones. These principles actually provide the language for aid workers to explain what they are doing and justify why they are doing it. This is routinely demanded in negotiations with the various stakeholders who can block a humanitarian organization from providing aid. The principles help convince those actors that their own objectives – whatever they are – will not be undermined by allowing humanitarian assistance to be delivered.

Independence implies political, financial and, perhaps most crucially, decision-making autonomy from any other actor who may be pursuing their own *de facto* authority. The suspicion that the agency may, in fact, be pursuing some other agenda will simply be too great.

Despite this practical application, the value of independence has been challenged. In the 1990s, a new point of view grew in direct response to the observation that humanitarian aid was often not alleviating suffering sufficiently, and sometimes was even creating conditions for prolonging it. According to this perspective, such perverse outcomes were because humanitarianism was unwilling to endorse any larger agenda – i.e. precisely because of its search for independence. This new doctrine held instead that aid should serve longer term efforts to resolve the underlying causes of suffering and so ultimately end the need for humanitarian assistance. In this way, humanitarianism becomes closely tied to laudable – but no objectives in the area where humanitarian action is implemented. This can be an extremely useful tool for negotiating this access in conflict zones. If autonomy is absent – or is merely perceived to be absent – then that humanitarian agency will probably be blocked from entering a given area – objectives like peacebuilding, development and often human rights.

This has important implications in situations like conflict settings where the legitimacy or motives of those pushing development, human rights or peace is precisely what is being resisted. As attacks on aid workers in Afghanistan and Iraq over the last fifteen or twenty years show, it can make humanitarian organizations look like agents of political agendas, and so make them targets for anyone resisting those agendas.

We might characterize this as a divide between a more ambitious humanitarianism that tries to resolve the causes of suffering but can only deliver it in areas controlled by actors who support their particular political approaches to this resolution, and a more humble humanitarianism that is limited to maintaining life in a more basic sense but in principle maintains the possibility of negotiating access to anyone in need of emergency assistance.

This represents the first element of the context in which the challenges that new technologies pose for humanitarian independence must be understood.

The second contextual element concerns the two camps that exist within discussions around using digitized data in humanitarian work. On the one side are a group pushing the potential of digitized data to improve tracking of aid distributions, improve efficiency, cut costs, improve donor accountability and create more dignified registration processes.

On the other side are people arguing that digitized data represents a new threat to the people about whom it is being collected due to its sensitivity. While this applies in all situations, and is usually understood as an issue of privacy, in conflict settings especially the difference between privacy and security becomes hazy, and failures to maintain privacy can carry imminent physical risk.

I want to suggest that these two camps represent something more than a simple division of advocates stressing the potential of the new versus sceptics emphasizing its dangers. The benefits and the risks of using cyber-based technologies that enable this digitized data production in humanitarian work may be linked, and indeed exist in direct proportion to each other. That is to say, the greater the benefits digitized technology brings, the greater the risks it brings too.

In military scholarship, there is increasing recognition of this possibility that cyber technologies create a so-called "capability-vulnerability paradox." For instance, the connection of real-world weapons systems via cyberspace allows for intensive coordination between weapons and automation of the resulting system, potentially delivering substantial gains in effectiveness. However, that increased presence in cyberspace expands the attack surface available for opposition cyber weapons to target, and the connectivity of these weapons systems means that any successful cyberattack could compromise the entire weapons array. As formidable as a cyber-connected, fully integrated weapons system may be, the very means through which it is created introduce new vulnerabilities that allow for it to be completely disabled.

Theoretically, cyber-connected humanitarian operations too could be similarly disabled. This is worrying, particularly in a moment when attacking humanitarian installations appears to be becoming more common. But it's not my focus here. Instead, short of completely disabling a system, I am interested in how greater use of cyberspace also facilitates the unauthorized access of stored data.

Deliberate, unauthorized access of data produced by humanitarians appears far from easily fixed. As a result of several new technologies now being deployed in humanitarian work, humanitarians are producing more and more detailed data about their environment, their activities and the people they assist. That data – both due to its quantity and its detail– is arguably more likely to be strategically interesting to parties in a conflict than was true in the past. In recent years, the media has reported several instances of humanitarian groups including *Médecins du Monde*, UNICEF,

the WHO and a host of organizations working in and around Syria, being hacked and their data being accessed.

In contrast to the consensus within IHL regarding the prohibition of cyberattacks that affect the real-world ability of humanitarians to provide aid, as long as a humanitarian group's operations are unaffected, the theft of its digital data appears IHL compliant. This is due to an ambiguity over what constitutes an "attack" in cyberspace.

More than this, some have argued that IHL provisions may actually create a duty to hack humanitarian groups. As we're all familiar with, IHL calls on belligerents to use the minimum level of military force required to achieve their strategic objectives and thus limit the risk of harming non-combatants. As a result, parties to a conflict should gather as much intelligence as they can from wherever they can so that their military actions can minimise overall suffering and loss of life. If data or information held by humanitarian groups operating in the theatre of war can assist in that regard – and new technologies are making this increasingly likely – then that information must presumably be accessed.

Clear agreements characterising such hacking operations against humanitarian groups as violations of IHL could help here. But this alone is likely to be insufficient. Attribution of attacks is simply too difficult. And in any case, we appear to live in a moment in which IHL is sometimes openly violated when doing so produces short-term strategic gains.

As such, maintaining independence depends largely on humanitarian groups resisting unauthorised access to their digital data. Two arguments underscore the difficulty they face achieving this.

First, quite simply the level of cyber security required is likely to be prohibitively expensive for humanitarian groups that are already dramatically underfunded.

Second, even if there is the money, the level of data protection required may simply not be possible for aid groups given the particular position they occupy. The increasing penetration of cyber space into so many sectors of political, economic and social life make it more and more difficult to isolate them from each other. In this hyper-connected environment, most cyber security experts do not believe it is feasible for an organisation to build an impenetrable system. Instead of impenetrability, cyber security doctrine instead aims for resilience. This accepts the possibility of unauthorised access and strategies accordingly. This reframes the issue as a probabilistic one; it is about reducing the chance of a breach to an organisationally acceptable level and preparing contingency plans to minimise the damage done if a breach occurs.

For states, militaries, businesses and private citizens, accepting the possibility of such breaches as a price to pay for the benefits of connecting to this cyber domain appears acceptable.

For humanitarians, this cost-benefit calculation looks different. The costs of losing data fall less on themselves than onto those they are seeking to assist, and – crucially for my argument – on the political and military actors whom they must convince to allow them in. As such, to the extent that humanitarian groups rely on perceptions of their independence to negotiate access, the organisationally acceptable level of cyber-intrusion risk must be close to zero. Why would a political or military leader allow a humanitarian organisation to operate in his or her territory if that agency cannot prevent itself becoming a vector for that leader's opponents to access data from which they can glean strategically valuable intelligence?

For humanitarian groups that use independence as a negotiating tool to gain access to some areas, this has implications on risk mitigation. It means a reappraisal of decisions regarding whether or not the risk of producing and storing certain data is justified. That calculation should not only be made according to the potential harm that data could cause if control of it is lost, although this of course remains extremely important. It should also take account of how strategically useful it may be to other actors pursuing different agendas. If data is too useful, it may represent too much of a target for those seeking unauthorised access.

What do we do if embracing new technologies improves the quality of the aid we can give, but significantly reduces the number of people we are able to give it to?

This is a new challenge I think cyber-based technologies, in particular, are confronting humanitarians with. And fundamentally it may not be a technical challenge concerning best practices, but one of ethics.

VIII. The way forward? A conversation on contemporary initiatives to address the new technology in warfare

IN THIS SESSION, EXPERTS ENGAGE IN A CONVERSATION TRYING TO FIGURE OUT THE POSSIBLE DEVELOPMENTS IN THE FIELD OF NEW TECHNOLOGIES IN WARFARE, ALSO IN THE VIEW OF THE EXISTING DIFFICULTIES IN INTERGOVERNMENTAL PROCESSES[*]

Chair:
Cordula DROEGE
Chief Legal Officer, ICRC

Kaja CIGLIC
Director, Digital Diplomacy, Microsoft

Thomas HAJNOCZI
Ambassador, Director for Disarmament, Arms Control and Non-Proliferation, Federal Ministry for Europe, Integration and Foreign Affairs, Austria

Cordula DROEGE

This is the last session before the closing remarks of this Round Table. This is meant to be a conversation between two experts on the subject. Over the past two days we have heard about new technologies in warfare from many perspectives: cyber, artificial intelligence, lethal autonomous weapon systems and outer space. We have heard from technical experts and thank you for trying to give the rest of us who are not very good at understanding these subjects a good understanding of and a good introduction to each of these technologies. We have heard about the potential human cost and, as a humanitarian organization, namely the ICRC, this is a particularly important issue for us. We have heard about legal challenges and I think one of the recurring themes that we have heard about refers first and foremost to how international humanitarian law (IHL) applies to these new technologies. There are a lot of panelists as well as people in the room who have reaffirmed this and have made a point of emphasizing it. We have also heard though that these technologies are today's technologies, but they are also tomorrow's technologies and the full

[*] The following discussion, based on the transcript of the recorded session, reflects the debate among the panelists. It has not been revised by them and does not commit them with regard to the views expressed.

capacities they have or will have, together with the risks they carry, are still in a way unknown to us. So, we are speculating a bit about the risks and that, of course, poses legal challenges as well.

Is the application of IHL a matter of interpretation? How do we interpret IHL in these new domains with these new technologies? Is it just a question of interpreting the laws in light of new technologies? Are there any gaps? We have touched on the way forward and I think the idea of this panel then is now to dig a little bit deeper into the way forward. I want to quote what Camille Faure said yesterday that, while it would probably be over ambitious to get complete consensus about the meaning of all the rules, there is still a sense that it is very, very important to have more and more clarity regarding legal understandings and as much common understanding as we can about how rules apply. For example, the question as to whether a cyber-attack was an attack under IHL came up. Do we need human control over the lethal autonomous weapons systems? If so, what does that mean exactly?

In the international sphere there are many initiatives that go in the direction of further clarification, perhaps new laws, perhaps new interpretations. Moreover, there are certain initiatives that concern all the weapon technologies we have heard about. There are UN processes on cyber warfare, for example, there is a group of governmental experts and an open-ended working group in the UN dealing with cyber questions. There is also a group of governmental experts on CCW dealing with lethal autonomous weapons systems. You have or had a group of governmental experts on outer space. But you also have many initiatives coming from the tech sector: Microsoft, and we will hear about this in a minute from Kaja, called for a digital Geneva Conventions a few years back. What does this mean? Last year, at the Paris Peace Conference, the Paris Call for Trust and Security in Cyber Space was adopted. Microsoft has also joined some partners to set up a Cyber Peace Institute. We also have self-regulation in the sector of artificial intelligence on the part of multi-shareholders such as Amazon, Google, Facebook, Microsoft, IBM. Google adopted AI principles in June 2019 out of a sense of urgency to have a human element in these new technologies.

It is, therefore, a sector in which the need for a way forward is felt very keenly by States, civil society and by the tech sector. All of us are struggling a bit to know what the best way forward is, also given the difficulties that exist, of course, in intergovernmental processes.

We have two people today on the panel who are really well placed to speak about these issues. On my right, I have Kaja Ciglic who has been one

of the leaders of Microsoft's significant work on protecting cyber space. She has been engaging with States, other companies, civil societies as the Digital Director for Microsoft and she has achieved really important results in that sphere. On my left, I have Ambassador Hajnoczi who is Head of Disarmament at the Austrian Ministry of Foreign Affairs and has been involved in the group of governmental experts on lethal autonomous weapons. I think both of them will give us really good insights into processes. We were meant to have a third speaker, Ambassador De Aguiar Patriota, who was chairing the Group of Governmental Experts on Outer Space and the one now on Cyber. Unfortunately, he couldn't make it. However, our two panelists will be able to address these issues.

So, I will perhaps begin with you, Thomas, if I may, because you represent the Austrian Governmental Group of Experts on Lethal Autonomous Weapons and it would be a good start to hear about that process and your views on it. Thank you.

Thomas HAJNOCZI

Thank you so much but first of all I would like to thank the Institute and the ICRC for this invitation. It is a real honour and pleasure to take part in this prestigious Round Table and I promise you I will really give you my personal opinion so it will not be attributable to my Government. Even so, my opinion very often coincides with the official position, but I think I can be much clearer on this occasion in that way.

You mentioned the other GGEs (Group of Governmental Experts) but they are not the same. Those created by the UN are a limited group, something like 25. Governments are invited to take part, for example, to discuss issues such as outer space.

However, the GGE LAWS is a group that was established under the Convention of Certain Conventional Weapons which implies that every High Contracting Party to this Convention can take part together with civil society, representatives from the tech sector, universities and so on. This GGE has now been going on for about five years. It works under the rules of procedure whereby everything has to find consensus. As a consequence, everyone has a veto which, as you can imagine, doesn't foster quick progress or very clear results. However, I think these discussions are very important because we are gradually building a better understanding of what the issues are.

Basically, this year, it officially adopted eleven guiding principles. Ten of them were already negotiated last year so it is not such a big step forward. Certainly, there are mixed views on the way as to how to proceed

further. It is clear that there will be a working group next year, however, it is still not clear how many days the group will be together – that will be decided in November by the meeting of the High Contracting Parties of the CCW because there was no consensus. One delegation wants to have less than all the other delegations. Only when you have consensus will you have a solution.

There is another bracket on the mandate. We were sitting there in August to elaborate a report and the good news is that at 3.15 am we found consensus on the text so you can imagine it was quite an animated night. We have the review conference in 2021 and the idea is that we wouldn't have to struggle next year for so many days. So, we have a 2-year mandate but what is at the same time very important is that we have to report on the progress of the next year. You know how it is when you have two years to do something - usually in the first year you do no hours and certainly a number of delegations did not want to see this happening.

So, the real problem is what should be the result of the work? There is a clear majority that shares the conviction on that. We need a legal norm because, as was said by Helen very vividly on the first day, norms are addressed to human beings and I think Professor Dinstein made a very persuasive find – when there are fake technological changes then you have perhaps to adapt the law and that usually takes quite a long time. And that's the problem here as we don't have time. The idea is that we would have the necessary prohibition of certain things before these new weapons are developed. That's possible – we did it in the CCW. And what is the fear of many, when we discuss guiding principles and so on which have a value and we are certainly positive to this, but that is somehow over taken by events and logically speaking it would make sense for all far progressed countries in this technology development to have a kind of clear idea as to how far we can go.

We're doing very interesting work. There is, of course, slow progress and we will see what will happen in 2021. It is quite clear that all the GGEs usually prepare the ground for a new legal instrument then it becomes an open-ended working group and it is a limited one – we could debate as to whether we need this. There is a proposal for a mandate and for a legal norm and a proposal is very simple – everything that is not under meaningful human control and critical functions would be prohibited basically and that is precisely the wording the German Foreign Minister used - the right line is of course the human control, and as the UN Secretary-General Guterres says: Machines with power and discretion to take lives without human involvement are politically unacceptable, morally

repugnant and should be prohibited by international law. And I think it's quite a strong statement for the UN Secretary-General, but it brings across in addition to the legal perspective, the ethical perspective. So, I would like to stop here.

Cordula DROEGE

Thank you very much Thomas. Perhaps before I go to a follow up question I already have I shall first give the floor to Kaja to talk about the initiatives you have been following very closely. Perhaps if you talk about the initiatives by Microsoft you could follow and talk a little bit about the other initiatives on cyber that exist in the UN system. Thank you.

Kaja CIGLIC

I wanted to say thank you as well.

I think for industries the challenge has tended to be - we often speak different languages and I'm not a lawyer. The question is if a cyber attack is made by a tech company it is often very different to a humanitarian lawyer. So, we had a small break down session with ICRC maybe a couple of years back to talk about the digital Geneva Conventions which lasted probably about 3 hours examining the text in detail. We speak very different terms, which is why a lot of the time there are misunderstandings. A lot of the initiatives industries are currently engaging whether they regard AI you mentioned earlier, and I can elaborate of them, exist outside the world of warfare and issues like that. A lot of initiatives are geared to deal with everyday problems and we should have principles that should guide us when developing technologies. We see a misuse of these technologies not just in times of war but also in times of peace and this is where the Microsoft focus on digital Geneva Conventions comes from and what drives a lot of our concerns.

Regarding the Microsoft proposal concerning the digital Geneva Conventions, which was something that we floated 3 years ago, this was a proposal to try and shake up the UN system more than anything else. We felt the UN processes had been ongoing since 1998 when the Russians proposed a first resolution on this issue and there's not been much progress in that period. I think we are now the 6th GGE – 3 came to a consensus report. Some of the consensus was really important - the 2013 one confirmed that States agreed that international law and human rights applied to cyber. Then the 2015 one highlighted 11 norms of behavior, for example, do not attack critical infrastructure, do not undermine the supply

chain integrity in cyber space, do not attack the entities that help countries clean up after an attack happens.

But then there was no report from the next GGE and everything seems to have stalled. So, we thought we should do something. Things shouldn't just stop. We were afraid as we had seen a few countries put forward proposals that would undermine the idea that international law applies, for example, and put a lot of emphasis on sovereignty alone.

So, the digital Geneva Convention proposal at a very high level calls for 3 things basically: greater accountability for State action in cyber space, particularly in times of peace; greater responsibility of the tech sector to actually engage more and understand that they provide services but they also need to ensure they are secure and need to be actively engaged in, for example, clean up and investment.

We reaffirmed and proposed a set of additional norms to the ones of the UNGGEs. One focused on election interference in cyberspace, another was reaffirming the supply chain making a point that when you look up a product there is a mass market – please don't interfere with those because they are not targeted – basically, when you attack, and you then attack a large proportion of the population. An example was the WannaCry Attacks a couple of years ago where one of the old Microsoft systems was exploited. It spread over night to millions of computers. Even though the potential target was two particular governments, countries from China to Denmark to the UK, and individuals were affected.

We went out after the WannaCry Attack and tried to find the victims – one of the most famous examples was when all the UK hospitals couldn't function for a couple of days, so surgery got postponed – it's not necessarily a financial cost but it's a massive personal cost to the person involved who say needed heart surgery only to be told that it's been cancelled and nobody knows why. We forget how much we are reliant on technology on the day to day basis and that an attack doesn't seem to be that important or doesn't cause death and destruction, but it can cause significant damage outside the war environment as well.

Following from that, Microsoft also launched the cyber security techno cord where at the beginning about 30 companies (now over 100) signed up to 4 principles. Two of them state that we will partner with each other, with civil society and others to improve security. The strong point there is that there was a lot of investment in proprietary technology like Microsoft invests a billion dollars a year in security alone but there is a lot of open source, a lot of legacy technologies out there but nobody really looks at

them so finding ways to get the industries and others to partner to make sure that they also get updated, fixed - all these things we wanted to do.

Another principle is focusing on empowering users and consumers to be aware of the security risk factors involved. A lot of attacks are successful because people just click on the link. So, unless you are aware, people will always click on a link so awareness raising is an important point.

Then the first 2 principles focus on a commitment not to support governments when they conduct offensive operations and to actively oppose exploitation of technology. So, whereas a government can come to a company and ask if a back door can be built, or a process in the product that can be exploited later on, the companies that have signed up say no. Basically, the first step is a general commitment to security investment.

Following from that I think that the Microsoft techno cord resulted in 65 governments signing the Paris Call for trust and stability in cyber space last year which is a first multi-stakeholder effort bringing together civil society, industry and governments on security. It's a 2-page document but we hope it signals a further commitment to some of these principles and norms. None of the principles in the Paris Call are new. They built on the norms that were adopted by the UN. They also build on civil society initiatives, for example, there is a global commission on security and trust in cyber space that put forward 6 norms. There's the industry initiative, there's the cyber trust techno cord. Siemen has a sort of Charter of trust really focusing on securing industry on an industrial scale rather than techno companies necessarily.

So, Paris Call brings them together in 9 principles and we are hopeful that this year – it's a first ever- even more countries sign. It includes everyone in the EU, all NATO countries except Turkey, the US, Australia. We are hopeful that more and more emerging economies will sign on as well.

Now we are at the point where we are starting the UN process again. We have the UNGGE discussion based on the resolutions sponsored by the US. They will continue to look at international law, confidence-building measures, and norms. It's a very traditional GGE, it has 25 member States. We kick started it and we're excited that this time around they have started to consult far more broadly to include cyber and industry as well – I think that is a new thing in this space. And then the other process is open-ended working groups sponsored by a resolution by Russia. That one is open to all member States and it has a mandate that really focusses much more on development but I think they will also try and understand how they can build on some of the work from the previous UNGGEs. They are starting

that work next week. I think their first meeting starts on Monday and I think there are 6 meetings. They only really have 1 consultation scheduled with a broader group of stakeholders in December and I think I would say that the point of how do we make this area successful, figuring out ways to get civil society and industry techno experts to work more closely together as technology evolves so quickly, would be very helpful.

Cordula DROEGE

Thank you very much Kaja for this first introduction and I'll go immediately back to Thomas to react to what Kaja was saying about all these attempts at regulation of cyber space which I think for some of us can be a little confusing as well because, as you mentioned, Kaja, there's this overlap between armed conflict and everything else. And there are many, many principles and attempts. How do you see that, from a government perspective, when on the one hand you have the UN processes and on the other you have private industry starting to push and perhaps to spearhead principles? How do governments react to this?

Thomas HAJNOCZI

Yes, thank you so much and of course to Kaja who made an excellent presentation.

It's a little bit confusing for us that at the same time we have GGEs and open-ended working groups.

It's not the normal way we work. We did have a GGE for a number of years but the last time it didn't come to a consensus on the results. Sometimes, when you move from a GGE to an open-ended working group some States say that it is not their moment, we didn't achieve consensus now so we should open up for everyone. And then, on the other hand, you have some States that need further work on this restricted framework before we can open up.

So, these two different views materialized into competing draft resolutions during the last General Assembly and the question was who was going to blink first? And there was a high expectation and hope amongst many countries that they would get together and come up with a common text which had been usual a number of years before, but no one blinked first. So, now we have two resolutions adopted by the General Assembly creating these two parallel bodies. That is a certain challenge of course, especially for the Chairs, but also for all the others because we wouldn't like to see a duplication of the work. And, as Kaja has pointed out, there are some ideas as to how we can have delineation of the work.

Whether this will really become a reality is something that we are all going to find out in the future.

Basically, I think there's a situation where, first of all, internet was not invented or created by governments. I think all of us know that internet was the invention of tech people and governments got into this game rather late – a little bit for the addresses, and so on. The US Trade Department was involved which has been changed in the meantime because there shouldn't be a State monopoly – their function was rather limited – they did it in a very neutral way.

So, that makes it a little bit different from quite a number of other things because, for example, development weapons and so on always have governments behind them. So, in the internet sphere we have this multi-stakeholder approach.

This is very important. In my view, it cannot be that governments themselves decide. You certainly have to involve not only the tech community but NGOs, academia, and so on. And when you go to internet meetings you will see all these groups there.

Then the second factor for me is – I mean, here we are speaking about cyber security but when I was Ambassador to the United Nations in Geneva one of the top issues was of course human rights in the digital age and we were active in this – there's a lot of human rights issues involved. So, basically, my own belief and that of many other people and governments is we would want to have a free internet – it should not be divided into national internets and someone is blocking what you can and what you cannot see.

This is very much engrained in all these discussions about possible regulations. Of course, we don't want to see a "everyone for himself" approach because it wouldn't function that way either. But there is an ingrained resistance in many quarters against the regulation of the internet by treaties that would be elaborated by States and would give a particular power or monopoly to those States. So, this is something we have to bear in mind when we to talk about all these little issues.

I also want to underline what Kaja said – I think it's extremely important that our education comprises the behaviour, the rights, the dangers, for example, when you're surfing the internet.

Today, digital issues are taught in schools but sometimes, very much only from a technical point of view and I think teenagers cannot be expected to be aware of all the dangers, the data and internet protection and so on.

Some things can be improved along the lines that Kaja has pointed out.

Cordula DROEGE

Thank you very much. And if I can just have a very quick follow up question on the issue of cyber and more particularly about the armed conflict aspect. You were very clear when you were talking about lethal autonomous weapons for which your think there's a need for new norms and the prohibition of certain types of weapons. I would like to know your views about legal gaps or not regarding warfare.

Thomas HAJNOCZI

What seems to be clear to me is that we must have a serious State response. The last GGE did some successful work on this and from my point of view it would be worthwhile to look better into the implementation now – how can this be implemented? And when we do so then we may come to a shared view as to whether further legal norms are necessary or not and I also want to point out the work that is done by regional organisations such as the OSCE because it came up with very practical confidence-building measures. So, a kind of mechanism that tells us how to proceed, in case there is a suspicion that someone hacked into our network, which is happening every day. What we need is a more cooperative approach. So, to answer briefly, I have not come far enough to make a final judgement as to whether we need new forms but there is a lot of space to make some real progress to achieve responsible State behaviour then we will past the test.

Thanks very much. Kaja, would you like first to react to that and you also said you had a follow up question for Thomas, so please go ahead.

Kaja CIGLIC

I would agree with what Thomas has just said. I think the question of implementation is something we struggle with as a society at the moment. On Wednesday, the first presentation was on how many States have offensive capability? And I think the speaker mentioned 30 – I think that number is actually a lot higher and the concerns we have a lot of the time are that these are the technologies that are effectively and comparatively cheap so a lot of countries can get their hands on them. Not all countries have processes in place that basically balance the possible consequences with the use of such a tool. And often these tools are multi-use – you could use Word to break into someone's system – it doesn't have to be something completely set aside from what you and I use on a day to day basis when we work. We haven't really seen a lot of governments react when they see attacks. Only in the last couple of years we have seen particularly Western

governments start to coordinate which is very good, call out actions. We are concerned when these attacks are attributed to a particular actor. They don't necessarily go back, and this is bad because, this is a UN norm, they just say that it's this person. So, we would like to see more countries committing to this norm that we have agreed to, reaffirm it by calling it out as well as using broader interpretations of it. This is where we see a gap – a lot of the norms are fairly broadly interpreted or defined or broadly written and not necessarily defined which leaves gaps. Countries, if they want to, can still look at them and find ways of getting around them. So, finding ways of getting them narrowly scoped or more easily scoped and ensure that countries actually comply with them, I think this is something we would really like to see.

We obviously see the need for more norms both on the part of the tech sector and of the governments. Some of the ones like the Paris Call have specific norms that call on the techno community to comply with. We think that everyone needs to raise the bar level and not just one particular community.

My question more than anything was maybe about how you see industry cooperation. It would be interesting to see and hear how you work with civil society, with industries and groups that are not necessarily just government.

Thomas HAJNOCZI

Thank you very much for the question. Obviously, I personally like to work a lot both with civil society and the academia. First of all, my level of knowledge is very limited and I can only benefit from their input and also on certain subjects you need real experts who have the technical or whatever kind of knowledge. I'm a fan of inclusive processes. For me, it's quite sad that at a number of meetings the States give their speeches, speeches, speeches then at the very end you can listen to NGO statements or it could be somebody from the tech sector. First of all, it's boring for all of us because what would be good would be to be interactive, and the earlier tech people and NGO representatives can speak the better you can interact. It's like having many meetings in the fringes and so on in these sectors and don't forget international organisations such as the ICRC. A stakeholder usually likes to call the ICRC the guardian of IHL so it's always worthwhile to include them in these processes and I must say that the interventions of the ICRC, for example, in the GGE LAWS make an impact.

So, I think we can improve in our working method in order to give more weight to the international organisations such as the ICRC, academia, tech people and NGOs. Like I say, I am a fan of the wider stakeholder model because otherwise in today's world we don't get there.

Kaja CIGLIC

Part of the reason why I asked this question was – when you ask a question as to how governments react when industry has an opinion, I would say, probably badly. What we have learnt from the digital Geneva Conventions was that it is, as I referred to earlier, really hard for us to speak the right language and this is why with military and AI we have had a position in the form of a blog which eventually will be made public and we've been struggling with it for over a year now I think. Part of the reason is that we talk at conferences like this, we listen and we talk to civil society, we listen and talk to the Defence Ministry and listen. We've all offended someone. So, it's hard for the industry to be helpful and to try to address the challenges that the governments are trying to address because we are just living in our own little space. So that's why I was asking.

Thomas HAJNOCZI

Yes, I think it's a very important thing. Sometimes we live in silence and we speak in other languages and we are not good in understanding each other. I have a very high esteem for a number of colleagues who are part of whatever tech community or NGO who can bridge this by the way they speak. And also on our side as diplomats we should try to bridge it. When you cooperate with whatever, NGOs or industry, there is a particular challenge when you are a diplomat because when one NGO tells you this and the next NGO tells you the opposite, when there is no common opinion among the important tech companies it's a little bit hard for you as a diplomat to deal with this. Therefore, NGO coalitions regarding nuclear weapons or explosive weapons in populated areas and similar organisations in the tech industry are very useful because you have much more power when you have a common voice and that is something that facilitates both the interaction and elevates your influence a lot.

Finally, I see fluidity between sectors, academia, tech community and NGOs because some of the greatest experts on, for example, LAWS, are professors who became active because they came to the personal conclusion that there must be a limit. These people are really knowledgeable. And remember Google where quite a few employees called for a stop because they didn't want to work for certain projects. So, I think

the days are over where we can neatly define here are the tech people, here are the NGOs, because they are all people with our beliefs and convictions. And I think that one of the good things in the tech industry is that employees can be heard.

Cordula DROEGE

Thanks very much. I find it quite interesting and I think one of the things that comes out here in a way is the role of civil society, and perhaps there is a speed scale whereby governments have a certain speed, the tech industry is perhaps a little bit faster but perhaps the fastest is civil society which puts pressure on both and we heard earlier about the Microsoft employees being unhappy about the endeavors of the US Government. We heard the UK representative also talking about the policy position of the UK on human control being very much informed as well by public perception before perhaps having done all the ins and outs of all the legal analysis. So that is also a phenomenon that is quite interesting. Perhaps just a small question before I open the floor to the participants would be – the relationship between these initiatives and academic endeavors.

Tying back to what you were saying much at the beginning, Thomas, that intergovernmental processes move slowly – the CCW moves particularly slowly because it is tied by the rule of consensus although someone told me very recently that apparently this is a myth and we need to look at this more closely. We are also, I think very frankly, in the disarmament field, at least in the ICRC, a little bit worried more about the dismantling of the disarmament architecture than the construction of the disarmament architecture. Professor Gloria Gaggioli talked about the fact that academics, and sometimes tech experts and think tanks take things into their hands and we have all these manuals produced, for instance, the Tallinn Manual on cyber warfare, there is now a manual or perhaps two being produced on outer space. Do these manuals influence you? What is the relationship between such second best solutions and what could be the agreements among States? Can they ever replace agreements among States? I suppose my question is already my own answer but it would be interesting to hear from you.

Thomas HAJNOCZI

Thank you, Cordula. You have said something very important, that, of course, all these processes are not proceeding in a void. We have a certain political situation today in international affairs. A number of the pillars of the disarmament architecture are falling apart. We still don’t know when

INF (Intermediate range Nuclear Forces Treaty) will start, whether it will be prolonged. We had some news concerning the CTBT and so on so we must be aware we are not in the most cooperative international environment today. There's a little example last year which, in my view, was unnecessary when we ended up with two parallel resolutions – it's just a little demonstration. So, from my point of view, it would be much better for all these processes if there were a more benign international climate – that would be very helpful. Smaller States try to contribute but of course their means are quite limited. The NGOs and the tech community, and certainly international organizations like the ICRC are working for this. These are the parameters in which our work is being undertaken.

Cordula DROEGE

Kaja, will you wait for governments?

Kaja CIGLIC

We have to at some point. I think it's the government's responsibility. I do see that is why we have the proliferation of certain initiatives. Why do you see the global commission putting through norms, why do you see Microsoft putting through norms? Because people realise that the geopolitics is not ideal. They are not sure that in the long run cyber is getting any better as more countries get weapons. So, we agree that we should act and we can't wait another twenty years.

But to your other question on whether academic conversation and things like the Tallinn Manual have their influence – I definitely think that they influence Microsoft. Brad Smith, our president, has a copy of the Tallinn Manual and a copy of the digital Geneva Conventions in his office and he has consulted them a lot over the past two years. We regularly consult with academics in this space on AI and human rights aspects as Google has an advisory committee. There are obviously teething problems with working with some of them. Very prominent human rights and experts make sure that it is interpreted on a line with the companies. I think you're right about customer pressure making sure that the company is going in that direction because it fears losing market shares but they also see a responsibility in this space.

I think there is a sub section of AI but last year we called for a regulation on facial recognition. The way we put it was that the race to the bottom was not just something that should happen but, basically, we need governments to set the minimum standards. We put forward some ideas and

we lobby for them but it's ultimately up to the governments to decide but companies can put pressure.

Cordula DROEGE

Thank you for this initial round of conversation. We have about half an hour left so I would like to open the floor to questions from some of the participants. Perhaps I shall take a few questions and then give the floor back to you and then another round.

Questions from the audience*

I have a question for the Ambassador. There are already a few articles that analyse the potential impact of Artificial Intelligence on nuclear weapons so my question is: Do you think that the current debate around Artificial Intelligence will affect in a certain way the upcoming talks on the NPT review conference?

I have more of a comment than a question. First of all, I think it is very interesting we are discussing these kinds of issues at these kinds of forums particularly where there is a combination of academics and State lawyers present. As a GGE member it would be helpful to provide some comments on behalf of The Netherlands. Since, as has already been pointed out, we are preparing for discussions at the UN level, I think it's important to highlight the background of what we are faced with as government lawyers preparing for this. Of course, we start from the basis of the 2013 and 2015 GGE reports as already pointed out by the speakers. We start from the premise that existing international law applies. And this is, as we have seen with a lack of a consensus report in 2017, may not be as easy to uphold as we may have thought at first. So, for now, although I agree with my colleague from Microsoft that in the future we will be looking at additional regulations but for now time is not there and we have to see how we can interpret current existing international law in the cyber domain. For IHL this means that, at least for us, it is important to re-state that it applies to cyber space and to all the domains we are discussing and that it should go

* Paragraph in italics reports the transcribed questions that were posed to the panelists. Names were removed for privacy reasons.

beyond what has been included in previous reports and, although it may sound logical, it is not that easy to discuss in all these forums.

As for States taking a position, it might be good for you to know that, like other States, The Netherlands has been working on this in the past years and we have actually sent a letter to our Parliament detailing the interpretation and the application of international law in the cyber domain. This letter has very recently been translated and it will be posted online in the coming weeks and for those of you who are interested we will be circulating it to colleagues that we have the contact details of. And lastly, then I shall stop my intervention, a little thing about the manuals that have been talked about throughout these few days. Of course, the Tallinn Manual for us, as has been pointed out, is not the law but it is a very helpful piece of information and because it also includes information from State consultations and we look forward to the manuals that will be coming in the next few years on outer space to hopefully include similar information both academic and a State perspective so it can help us in determining our positions also on those issues.

Monsieur l'Ambassadeur, Excellence, vous avez évoqué ce phénomène de vivre en silo et ne pas parfois écouter les voix qui viennent d'ailleurs. Nous avons entendu depuis ces trois jours qu'il y a de la nouvelle forme de la guerre beaucoup plus urbaine, beaucoup plus de technologie avec des défis de fabrication entre le droit humanitaire international, le droit international des droits de l'homme, le droit des réfugiés, ecce. Est-ce que vous prenez en considération dans vos travaux actuels les obligations des Etats en matière de l'homme? Mais aussi les constatations, les observations et les recommandations des organes des traités et les organes d'experts comme les commissions d'enquête mises en place par le Conseil des Droits de l'Homme ?

Ma deuxième question concerne ce que nous avons entendu ce matin et tout à l'heure d'ailleurs que dans la phase d'observation il y a beaucoup plus d'implications de l'intelligence artificielle et beaucoup moins dans l'implication humaine. Lorsqu'il y a une erreur d'observation, une erreur d'analyse de l'information qui est responsable - la machine, l'algorithme ou la société privée qui l'a produite ou l'Etat ?

Ma troisième et dernière question et je serai très court. Est-ce qu'il y a des moyens de renforcer l'usage pacifique et positif de ces nouvelles technologies cybernétiques ? Je prends, par exemple, la Commission d'enquête Myanmar mise en place par le Conseil des droits de l'homme qui a utilisé des images satellites, a utilisé l'analyse produite par l'institut des

nations unies pour l'analyse des images satellites pour prouver des crimes de guerre et des crimes contre l'humanité en corroborant les témoignages, les vidéos avec les mouvements des troupes démontrés, documentés par ces nouvelles technologies.

Cordula DROEGE

Before I open the floor again I shall give it back to our two panelists. Perhaps first Thomas and then Kaja.

Thomas HAJNOCZI

Thank you for your excellent questions. Starting with the first question concerning the impact of IA and in general of cyber issues on nuclear weapons - that is a very important point and it is considered, at least by us. I had the privilege to moderate a panel during the NPT legal conference precisely on a report written by NTI that is the Nuclear Threat Initiative, which is a Washington based think tank, but sponsored by Senator Sam Nunn and the former British Defence Minister, Des Browne. What does it mean for the reliability of nuclear weapons - that hacking and those things are possible? Actually, I found findings in the report quite shocking because the base line was due to the possibility that nuclear deterrents were not reliable anymore. And that was written by Sam Nunn and Des Browne who are certainly very respected, and, depending on who is in government, very close to the US Government. So that is a major concern for all of us who deal with nuclear weapons. These findings show that it would be good to have a new look at some of these paradigms that have not been questioned for many years because our world is changing as should our approach to nuclear weapons in order to be valid.

With regard to my colleague from the Dutch Ministry of Foreign Affairs, personally, I must say that I pretty much go along with what you discretely alluded to. For someone who has studied law some forty years I think it is sad that we have discussions on whether international law applies. How can we question that international law applies? International law certainly includes international humanitarian law and human rights law. Therefore, for example, we always make sure that this is clearly stated and people who are part of the negotiations in the GGE LAWS know that. So, I think on the one hand it's good that we have one of these common outcomes and on the other hand, frankly speaking, it's impossible that international law does not apply because it is not up to States to say if the law applies or does not apply. It's the nature of any legal framework that it applies.

Turning now to the member of the UN Committee against Torture, yes, human rights are important in that context and having worked on human rights issues and having had the pleasure to have contact with the treaty bodies, I know how much expertise there resides. I think we should all try to include their work more in our analysis. I know that many foreign ministries tend to say that now I am on the disarmament side, I don't hear anything on human rights. I just shut my eyes and go for the disarmament files. I don't like this. I don't think it's helpful. Certainly, it is easier in a smaller structure like the Austrian Foreign Ministry. Michael Trenton, Former deputy, is the head of the human rights section. We often have lunch together, and he sends me some things he thinks would be interesting for me and I send him things I think would be interesting for him. So, I think we should try and fit this in. Whether we do enough I would clearly say no but we should strive to get better.

Then there is the issue as to who is responsible. There's this accountability/responsibility problem, of course, but once machines do certain things pretty much on their own there is always this discussion on who should be held accountable. Again, it's what Helen said: the law is directed to human beings and not to machines and it shows once more the necessity to have a human control. In the armed forces it must be clear who is in command and who is responsible.

It is certainly very important that we do not fall into the trap of being over critical of the main advantages that new technologies, like AI, bring to us. I think there is a lot we could bring but at the same time I do not think that everything that becomes possible should be done. There must be an ethical and legal line but you don't get there so quickly. Technologies can help in many ways, I would say, certainly in the peaceful area but even, when you have more precision, as is the case of certain arms we discussed this morning, in urban warfare it can help if you only hit this target and you don't risk the lives of hundreds of civilians so, there are somethings that can be done.

Kaja CIGLIC

I wanted to thank my colleague from the Dutch Government because they are really driving a lot of the work in this space from hosting the London process to creating the Global Forum for Cyber Security Expertise which is a forum that really focusses on capacity building in this space and I think that this is attuned to the last point as to how technology, apart from AI, can be very beneficial in a lot of these areas. And what we need for that is greater awareness, understanding of how to use it, of what's available

and ethical principles involved. You would think that the example you gave was a great example but there are even little things like automatic translation when you go into a country. I know it works better in some languages than others but we're getting there. AI is still learning. It is early days. I think, at the moment we all see it as a challenge, as this big thing that is coming. I think these are incremental steps. We have been using it for a while. Another example is, that AI actually improves cyber security a lot because we are able to get the computers to defend themselves from the millions of attacks Microsoft gets every week and this is definitely beneficial.

As far as human rights are concerned – part of my team is called the Human Rights Team – and their task is both to look outside at external commitments so Microsoft can make and sign up to and to ensure we work with the product teams and with the sales teams to make sure that we comply with those principles. As for facial recognition we have declined several law enforcement agencies. There are certain things we do not sell to certain governments. So, there is a level of attention in respect of those principles.

I have a question for Microsoft. I think it is really interesting because the law companies are reluctant to wait for regulations from governments telling them how they should be running their business, so it is an interesting approach that Microsoft is waiting for that to be developed. But my question in the meantime as we wait for that regulation is: what are the different measures that Microsoft and other companies are taking to reduce the attack surface? And the Geneva Conventions you are advocating is about protecting people and some of the tech companies are responsible from the devices to the software, to the broader infrastructures. So, as we wait for government regulations what are the other measures that tech industries can practically be taking without being asked by governments to do so?

Just two developments which have not been addressed by our panelists: one is the UN Secretary-General Digital Panel which was co-chaired by and Jack Ma and Melinda Gates that issued a report this June where work is ongoing to see how it would fit in with the UN Secretary-General's agenda for next year.

The second one is the Global Tech Panel which is a forum where there is an exchange between my High Representative, Ms Federica Mogherini, and the tech industry, including Microsoft Google and academics, that

discusses the principle of tech for good with a number of tracks being developed in close cooperation with members of the tech panel. We hope we will have an assessment somewhere nearing the end of the mandate of my High Representative on the 1st November to see if we can progress in that kind of innovative dialogue with tech in order to ensure that politics and tech can work together to address the issues including the issue of laws. The Ministries of Defence of the European Union were briefed on that last Wednesday at an informal dinner where tech members were present and we have a sincere hope that we are going to be able to progress that agenda. That's just a comment I wanted to make.

In this panel we have witnessed one debate on the application of international law in the cyber domain and, regarding this matter, there are also some people who have different reservations, different perspectives in this regard. But if you go inside the reports of the GGE, especially the report of 2015, we realise that there is one key concept regarding the responsible behavior of States. I do believe that the experience of the international community in IHL and this new space, we can take the main elements of the responsible behavior of the States. So, in this regard, I would like to ask Ambassador Hajnoczi who is involved in this matter and Ms Kaja as well: what are the different elements of this concept of the responsible behavior of the States?

Kaja CIGLIC

I forgot to highlight earlier the EU sanction regarding cyber – I think that is a very good initiative to start driving State accountability and State behaviour in cyber space. We were really excited when the EU adopted it and we are looking forward to it being applied in person.

On the points about what industries can be doing there are two things – state behavior – please don't attach important info structures – we wait for regulations, we call for regulations – it's hard for us to do very much about their behavior. On the other hand, concerning more traditional, national level regulations, for example, we called for the US air protection regulation in 2005. There are definitely principles and investments that companies can make and are making. I have spoken a little about the cyber security technocore) and the efforts that group is undertaking and the AI principles whether you look at the ones Google has adopted or Microsoft has adopted or another set of guiding documents that help decision-making easier in companies. Another investment area in the partnership with AI you mentioned earlier bringing together cross industry, cross academia

especially groups to work on these hard issues because I think we need to acknowledge, as you mentioned Thomas, that governments can't solve this alone, nor can industries.

The importance of actually funding and bringing those groups together to try and find solutions is critical. What we do as part of business is investment in business security. I mentioned earlier we invest over a billion dollars a year on security. We have over 3,000 people working on security. It is a sort of interesting game. People say that being on the offensive is often the best defense but we don't have that luxury so we need to invest in defenses. In our case, we do a lot to bring forward new technologies that are more secure, adopt standards, sometimes develop standards that we then share with industries. I would say that a lot of the challenges we see in these spaces, not necessarily with the bigger companies that are more mature in their space and have a lot more funding. I would never say that everyone is completely secure because no one is. It's more a question of risk management. But the issue is more that the incentives for new technologies is to be first to market and security is not what people really pay attention to at that stage and people don't pay for it. Looking at individuals we always want to buy products cheaply – we don't necessarily want to put a premium on security at this stage and I think that also is something that needs to change.

Thomas HAJNOCZI

I am very happy to hear the voice of the EU because certainly Austria is a EU member State and one that is very committed to European cooperation and I think that one of the big sins of this panel is that we didn't mention the UN GGE panel report. I think it was a very interesting undertaking to do this and there are very good recommendations in it and now, of course, we have to work on them.

There are some principles that are highlighted in the report such as inclusiveness, respect for human rights and dignity, diversity and so on, and human centredness - systems should be designed to maximize benefits to humans and to ensure that humans remain responsible for decisions. These are very good guidelines. There are other interesting proposals and some of these recommendations show that there is an attempt to find some common ground especially when considering the architecture and I am quite interested to see what the concrete follow up will be.

You mentioned the Global Tech Panel yourself, and there are a number of important initiatives on the way. The EU, in my view, can play and should play a very constructive and positive role in all these undertakings.

The EU parliament, for example, has adopted, with a surmised majority, a resolution on laws making it clear that nothing should exist that is not under human control. We have common EU statements during the GGE LAWS. Of course, there are some new answers amongst us but so far we are all for a normative framework, for example, and for new sanctions for cyber. Personally speaking, I think that was a very interesting and important step and the real difficulty that has been expressed by many before, is the retribution. This shows that the EU is serious about the responsible behaviour of States.

So my Iranian colleague referred to this concept and, for me, at the end of the day, the real test is how do we implement it? And I would like to see our work concentrating on this. I think this is the right concept but a concept is only really valuable when we also implement it.

Cordula DROEGE

We will have to close this very interesting panel because we now have some closing words from Judge Fausto Pocar and Dr Helen Durham. Join me please in thanking both our very interesting panelists for a lot of new insights and learning about all these different initiatives and principles that exist around new technologies. And, of course, our dear interpreters who as you all realise, have a particularly difficult time with this type of format of conversation, so a special thanks to you.

Concluding session

Closing remarks

Helen DURHAM
Director of International Law and Policy,
International Committee of the Red Cross

We have come to the closure of this event. This is always a sad time as I do believe that the annual Round Table is a unique place in the international law landscape where, as Professor Pocar said at the start, we come together with different backgrounds, different experiences, different skill sets. We are drawn in by the Sanremo spirit and genuinely discuss issues that we need to move forward on. I have greatly valued the interventions from all experts and the audience. I think the multi-disciplinary nature of the subject we have been grappling with has been marking the last few days.

We started with a keynote address marking the 70th anniversary of the four Geneva Conventions. What we heard resonated throughout the rest of our sessions: the law is not a static object. International law – particularly IHL – has to be flexible in some ways to be able to absorb changes. Still, as our keynote speaker highlighted, there are foundational principles that have served the test of time and that need to be upheld. As the ICRC, as the guardian of international humanitarian law with that particular international legal personality and a mandate agreed upon by States, we are very clear and very aware that international law is owned and developed by States. If we look into the fascinating history of the Red Cross and Red Crescent Movement, we see that the ICRC and the Movement have had an influence on the development of the foundational principles of IHL. In the ICRC's archives, we find the ICRC's proposals that were submitted to the International Conference of the Red Cross and the Red Crescent in Stockholm in 1948. And some of these proposals were included, a few years later, in the Geneva Conventions. We have the 33rd International Conference of the Red Cross and Red Crescent this year and we must not forget that this is another place where ideas are generated and expertise is gathered

After the keynote address, we started with three experts examining how international humanitarian law has reacted to changes in the nature of warfare in the past. Our experts did an excellent job in presenting a variety of options to address changes in warfare that may also inform our thinking on the way forward. These include the application and interpretation of

existing treaties and IHL principles; the development of new treaties; the development of the law through courts; and the publication of academic manuals.

We then turned to our first panel on new technologies, which focused on cyber warfare. The panelists underlined very well the real threat to humans and civilian infrastructure that the use of cyber operations in armed conflicts can pose; the challenges that derive from the rapid development of cyber technology, the spread of this technology, and the difficulty of attributing cyber attacks. It became also clear that while IHL applies to and restricts cyber warfare, we need more debate among States on how existing rules should be understood in cyberspace.

In the next session, we had an excellent debate among two experts on the use of autonomous weapon systems in armed conflicts - and I must admit that having the format of a moderated debate between speakers is a very good way to go. The examination of concrete scenarios helped us to dig deep into the legal and ethical issues we need to consider on the use and regulation of increasingly autonomous weapon systems. What does autonomy in warfare mean? What are we really talking about?

Closely related to autonomous weapons, the next panel examined the use of artificial intelligence and machine learning in warfare. It was extremely interesting to hear from a technology expert what AI and machine learning can do today, and, importantly, what they cannot do. Likewise, it was great to have a military practitioner sharing how States are currently using AI in military operations. And of course, we had two lawyers presenting limitations that IHL provides for the use of artificial intelligence. I think one important message that came both from panelists and from the audience was that information provided by machines, be it on targeting or detention, needs to be considered with significant caution – machines cannot be followed blindly.

In the last panel of the second day, we turned to outer space warfare, learning about States' outer space capacities and operations. It was fascinating to hear in-depth legal discussions on how outer space law, the United Nations Charter, and IHL regulate the use of force in outer space, and to learn about some of the ongoing debates in these fields. And here, I saw some analogies to debates that we also have with regard to other technologies, such as what constitutes a 'hostile act', or which operations would amount to 'attacks' as defined in IHL.

This morning, we then turned the application of new tech technologies in urban environments. What I recall from the discussion are valuable insights on how new technologies will allow gathering more information,

synthesize it and thereby allow commanders and soldiers to have better situational awareness and apply IHL. We also heard a very clear word of caution: new technologies are unlikely be the silver bullet to the various challenges of urban warfare.

After focusing mostly on military operations, we then turned to the prospects and challenges of using new technologies in humanitarian operations. In my view, the challenges related to innovation in the humanitarian sector warrant careful reflection. I would reiterate two points from our discussions: First, there is a great need for innovation in the humanitarian sector, but we should not solely focus on technology but on all ways in which we can assist affected populations. And second, if new technology is used, there are important data protection challenges. In fact, it appears that there is an important dilemma building up: the more data is gathered in operations, and the more systems are digitalized, the more they may become the target of hacking and be vulnerable to it.

And finally, we had a significant and multidisciplinary conversation – with an eminent diplomat, an expert from the tech industry, and the ICRC's Chief Legal Officer, on policy approaches going forward. I must say: much work lays ahead of us!

Until next year I wish that you all use and build upon the lessons learned, experiences shared and ideas discussed during this Round Table. I hope to see you all in 2020 to mark the 50th anniversary of the Institute and to continue our important conversations.

Closing words

Fausto POCAR
President, International Institute of Humanitarian Law (IIHL)

Thank you, Helen, for your closing words on the achievements of this Round Table, as well as and for your kind words addressed to me as President of the Institute. Let me say, in turn, that it has always been a pleasure, over these last years, to share with the ICRC as an institution the annual Round Table and, in particular, to have personally shared with you its concluding words, as well as the task of identifying, at the preparatory stage of each Round Table, in perfect tune with you, the issues we were going to address on each occasion and the way we were going to deal with them. We have always been a good and efficient team and I am grateful to you for the high level of friendly partnership we were able to reach in the interest not only of our respective institutions, but, even more, of humanitarian law and of the human beings who benefit from its progressive implementation.

At this late hour, I will not take more time to go through the programme and comment over all the issues we discussed in the sessions of the Round Table. How IHL responds to the challenges of new technologies has been and will always be a theme for provisional conclusions, especially because new technologies develop quickly and bring about new challenges for the full application of the basic principles of IHL – distinction, proportionality, precaution and humanity – which must always be at the core of our attention. One point which was made at the beginning of the Round Table was that, whatever the view held on the need for adaptation of the law to new technologies, the basic principles of IHL remain valid and must be respected. Consequently, should new technologies not allow for their respect, they would entail violations of international law.

Indeed, on the one hand, technologies may show advantages in the application of the said basic principles, but, on the other hand, they may entail disadvantages, shortcomings and problems. We also heard today that technology may contribute positively to the missions for the assessment of violations of IHL and may assist in that assessment, even in courts. However, new technology or parallel technology may also contribute to hiding violations and make their assessment more difficult, so that accountability for crimes might become less easy to establish in a court of law, be it an international court or domestic jurisdiction.

But it is not the right moment now to go into details of the Round Table and to discuss again to what extent IHL may succeed in facing all the challenges raised by new technologies. It is also doubtful that a discussion on the role of the human resolves the problem. The problem is that, while it is difficult to ensure that new technologies respect the afore-mentioned principles, it is even more difficult to make sure that humans behave correctly and prevent them from making mistakes, intentionally or unintentionally, both in the application of traditional and new technology.

Education and training in IHL, as national military academies are mandated to carry out, are essential in this regard. However, I wish to stress that, if lessons can be drawn from the use of new technologies in modern warfare, where military operational decisions may be taken far from the actual battlefield, or perhaps where the battlefield has become global, it has become more and more imperative that dissemination and education on IHL reaches not only the military, but any person involved in armed conflicts, including at the political level, and, especially in light of frequent non-international conflicts, and the population at large.

The Sanremo Institute has been at the forefront in the dissemination and training of IHL since its foundation. Thousands of persons, military and non-military, have participated it its activities, in its courses, in its round tables, and have been trained using its manuals elaborated by competent and independent academics as well as operative international experts. The Sanremo Institute will continue to play a significant role in this regard in the years to come, and the annual Round Table will remain an important gathering of experts where the developments of IHL and its implementation can be discussed.

This Round Table has been very productive in creating a fruitful exchange of ideas through thorough and competent discussions. I am confident that it will be the same in the future. Next year will be a turning point for the Institute, as it will celebrate its 50th anniversary. The next round table will certainly be devoted to an assessment of the achievements reached in half a century of active life, and at the same time it should look to the future and discuss where IHL is going and where the focus of our endeavours should be placed. We will try to organize a memorable round table for the 50th anniversary and I hope to welcome you all in Sanremo again, even if I will do it in a different capacity.

Once more, I would like to express my gratitude to the panelists and the moderators who conducted the various sessions. Admittedly, it was not easy to find the most appropriate speakers, mostly because experts in technology are very busy and were not easily available to join us, and I

wish to thank all of them for having come to Sanremo in these days to give their excellent contribution to the Round Table. Special thanks go to all those who participated in the debate and to the interpreters for their important contribution. Finally, let me renew my thanks to the staff of the Institute and of the ICRC whose commitment has made this Round Table a great success; I am deeply grateful for the efforts you have made.

It remains for me to wish you all a safe journey back to your homes in the hope of seeing all of you here again in Sanremo next year.

Acronyms

AI	Artificial Intelligence
AMRAAM	Advanced Medium-Range Air-to-Air Missile
AP I	Additional Protocol I
AP II	Additional Protocol II
Army JAG Corps	Army Judge Advocate General Corps
ASW	Anti-Submarine Warfare
C3	Consult, Command, Control
CA 3	Common Article 3
CCW Convention	Convention on Certain Conventional Weapons
CIL	Customary International Law
CIWS	Close-In Weapons System
CNES	Centre National d'Études Spatiales
COPUOS	Committee on the Peaceful Uses of Outer Space
CPS	Cyber-Physical Systems
CPU	Central Processing Unit
C-RAM	Counter Rocket, Artillery and Mortar
CTBT	Comprehensive Nuclear-Test-Ban Treaty
DARPA	Defence Advanced Research Projects Agency
DCDC	Development, Concepts and Doctrine Centre (UK)
DLT	Distributed Ledger Technology
DoD	Department of Defence
DPH	Direct Participation in Hostilities
DU	Depleted Uranium
ELDO	European Launcher Development Organisation
ESA	European Space Agency
ESRO	European Space Research Organisation
FFAO	Future Framework for Alliance Operations
GCs	Geneva Conventions
GGE	Group of Governmental Experts on Lethal Autonomous Weapons Systems

GMES	Global Monitoring for Environment and Security
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
HMI	Human Machine Interface
HQ	Headquarters
HQ ACT	Headquarters of the Allied Command Transformation
HRL	Human Rights Law
IAC	International Armed Conflict
ICBM	Intercontinental Ballistic Missile
ICC	International Criminal Court
ICCPR	International Convention on Civil and Political Rights
ICJ	International Court of Justice
ICL	International Criminal Law
ICRC	International Committee of the Red Cross
ICS	Industrial Control System
ICTR	International Criminal Tribunal for Rwanda
ICTY	International Criminal Tribunal for the former Yugoslavia
IFF	Identification Friend or Foe
IHL	International Humanitarian Law
INF Treaty	Intermediate-range Nuclear Forces Treaty
IO	International Organisation
IoT	Internet of Things
IP	Internet Protocol
ISR	Intelligence, Surveillance, Reconnaissance
IT	Information Technology
LEGAD	Legal Advisor
LOAC	Law of Armed Conflicts
LTMT	Long Term Military Transformation
MILAMOS	Manual on International Law applicable to Military Uses of Outer Space and the Protozone
MITM	Man-In-The-Middle
MoD	Ministry of Defence
MOUT	Military Operations in Urban Terrain

MUSIS	Multinational Space-based Imaging System
NATO	North Atlantic Treaty Organisation
NDPP	NATO Defence Planning Process
NGO	Non-Governmental Organisation
NIAC	Non-International Armed Conflict
NPT	Treaty on the Non-Proliferation of Nuclear Weapons
NSAG	Non-State Armed Group
NSL	National Security Law
NTI	Nuclear Treaty Initiative
OODA	Observe, Orient, Decide, Act
OTAN	Organisation du Traité de l'Atlantique-Nord
OTS	Orbital Test Satellite
OS	Operating System
PC	Personal Computer
PGM	Precision Guided Munition
PLC	Programmable Logic Controller
PNT	Positioning, Navigation, Timing
POC	Protection of Civilians
POW	Prisoner of War
ROE	Rules of Engagement
SATCEN	(EU) Satellite Centre
SATCOM	Telecommunication Satellite
SFA	Strategic Foresight Analysis
SIS	Safety Instrumental System
SOHO	Small and Home Office
SSA	Space Situational Awareness
SST	Space, Surveillance, Tracking
TIRA	Tracking and Imaging Radar
UNICEF	United Nations Children's Fund
UNOOSA	United Nations Office for Outer Space Affairs
WFP	World Food Programme
WHO	World Health Organisation
WWII	World War II

Acknowledgements

L'Istituto Internazionale di Diritto Umanitario ringrazia vivamente i Governi e gli Enti che hanno concesso un contributo finanziario o il patrocinio per la Tavola Rotonda.

The International Institute of Humanitarian Law warmly thanks those Governments and Organisations that have given either a financial contribution or their patronage on the occasion of this Round Table.

L'Institut International de Droit Humanitaire tient à remercier les gouvernements et les organisations qui ont accordé leur appui financier ou leur patronage à l'organisation de cette Table Ronde.

ARMÉE SUISSE

BRITISH RED CROSS

COMITÉ INTERNATIONAL DE LA CROIX-ROUGE

COMUNE DI SANREMO

CROCE ROSSA ITALIANA

CROIX-ROUGE MONÉGASQUE

DÉPARTEMENT FÉDÉRAL DES AFFAIRES ÉTRANGÈRES, SUISSE

MINISTERO DEGLI AFFARI ESTERI E
DELLA COOPERAZIONE INTERNAZIONALE

QATAR RED CRESCENT

Whither the Human in Armed Conflict? IHL Implications of New Technology in Warfare

This collection of contributions made by renowned international experts and practitioners explores the implications of international humanitarian law and the growing role of new technology in warfare.

The 42nd Round Table on current issues of international humanitarian law focused on some of the fundamental legal questions arising from the increasing military use of autonomous weapons and cyber technologies within military operations and the broader context of international security. Experts highlighted the threats posed by the conduct of cyber-attacks against civilians and civilian objectives, as well as those caused in conflict areas by the expanding use of unmanned weapon systems with reduced or no human control.

The Round Table provided a forum to discuss relevant topics related to the heightened development of technology with regard to military and security issues, including the applicability of IHL in outer space military operations and the role of new technologies in humanitarian operations

The **International Institute of Humanitarian Law** is an independent, non-profit humanitarian organization founded in 1970. Its headquarters are situated in Villa Ormond, Sanremo (Italy). Its main objective is the promotion and dissemination of international humanitarian law, human rights, refugee law and migration law. Thanks to its longstanding experience and its internationally acknowledged academic standards, the International Institute of Humanitarian Law is considered to be a centre of excellence and has developed close co-operation with the most important international organizations.

FrancoAngeli
La passione per le conoscenze

€ 35,00 (u)

ISBN 978-88-351-0883-2