Written evidence from International Committee of the Red Cross (ICRC) (TFP0029)

About the ICRC

1. The International Committee of the Red Cross (ICRC) is an impartial, neutral and independent organisation whose exclusively humanitarian mission is to protect the lives and dignity of victims of armed conflict and other situations of violence and to provide them with assistance. The ICRC also endeavours to prevent suffering by promoting and strengthening international humanitarian law (IHL) and universal humanitarian principles.

2. Established in 1863, the ICRC is at the origin of the Geneva Conventions and the International Red Cross and Red Crescent Movement, alongside the International Federation of the Red Cross (IFRC) and national societies around the world, including the British Red Cross.

3. The ICRC directs and coordinates the international activities conducted by the Movement in armed conflicts and other situations of violence.

"New developments in digital technologies are taking place at a startling pace, affecting the way we live, the way we work and even the way we think. They hold great promise for humanity." – Peter Maurer, ICRC President

Summary

1. The International Committee of the Red Cross welcomes this inquiry on a rapidly evolving, cross cutting area of industry, public policy and indeed all of our daily lives. The scope of this inquiry is especially wide given the many ways in which new technology can impact the relationship of the UK with the rest of the world. The Integrated Review (IR) made clear that science and technology would sit at the heart of not only foreign policy, but its defence and development positions also. We note the focus on technology's increasingly significant impact on and within conflict, and this is also a growing area of focus for the ICRC. However, where there is conflict, it follows that there is humanitarian need. Our efforts are dedicated to understanding and conveying the humanitarian perspective of new technology as it entails a new risk to civilians, which needs to be proactively identified and minimised.

2. The ICRC is continuing to develop and refine positions and has done so for decades e.g. in the realm of cyberwarfare[1]. These technologies, and our positions and concerns will be addressed sequentially, although an overarching concern through these distinct technological capabilities is the protection of civilian populations affected by conflict both on the ground and in the virtual space. A core message has emerged that, where artificial intelligence is concerned, a human centric approach must be the priority.

3. These technologies have arguably seen so rapid a development and prevalence in our lives, that the risk of exploitation and harm is almost impossible to identify, much less mitigate. As the call for evidence in this inquiry states, not only has the UK Government "traditionally

[1] Twenty years on: International humanitarian law and the protection of civilians against the effects of cyber operations during armed conflicts | International Review of the Red Cross (icrc.org)

been at the heart of the international system," it has an explicit intention to "be a world leader in shaping international rules and behavioural norms relating to new technologies[2]." The opportunity here for the UK to use both its influence and expertise to ensure a human-centric, and indeed humanitarian, perspective is at the forefront of global thinking is both timely and critical.

Cyber – Avoiding Civilian Harm

4. Cyber is an ever more encompassing term, but for the purposes of this submission, we have found it useful to break it down into two separate categories as can be applied to the conflict and humanitarian contexts. Firstly, we will address offensive cyber operations, followed by cyber security and the UK's stated intent to lead on capacity building in vulnerable contexts.

a) Offensive Cyber

5. The IR heavily references cyber and states that the UK is the third most powerful cyber nation in the world. The creation of the National Cyber Force (NCF) and the declaration of active offensive cyber operations predate the IR's publication. The UK's explicit declaration of strength identifies it as potential key advocate for the application of IHL in cyber warfare. Ahead of the publication of ICRC guidelines, 'Avoiding Civilian Harm in Cyber Operations' in June, our emphasis here is on understanding and reducing the *human cost* of such operations, whether they occur in the so called 'grey zone' or – as is the case with the UK's declared offensive cyber operations – firmly within the applicability of IHL.

6. Fortunately, the use of cyber operations by militaries have not led to dramatic humanitarian consequences to date. However, cyber-attacks against, e.g. transportation systems, electricity networks, or chemical plants are technically possible and could result in high numbers of civilian casualties and damage. Cyber operations could cause physical damage and affect the delivery of essential services to civilians, as recently highlighted in an expert meeting with military strategists[3], convened by the ICRC.[4]

7. Attacks using cyber capabilities in armed conflict are bound by the rules of IHL, This means principles of distinction, proportionality and precaution in attack and the special protection afforded to hospitals and objects indispensable to the survival of the civilian population are applicable. However, there is a real risk that cyber tools are not designed or used in compliance with IHL. For example, during armed conflicts, the employment of cyber tools that spread and cause damage indiscriminately is clearly prohibited. Yet, some cyber tools, designed and used to target and harm only specific objects, can cause indiscriminate harm due to the interconnectivity that characterises cyberspace. Due consideration should be given to the greatest extent possible to limit the unintentional (as opposed to deliberate repurposing) spread of these weapons.

[2] https://committees.parliament.uk/committee/78/foreign-affairs-committee/news/150600/new-inquiry-tech-and-the-future-of-uk-foreign-policy/
[3] Attendees included UK Ministry of Defence staff, allies and UK-based academics
[4] https://www.icrc.org/en/document/potential-human-cost-cyber-operations

8. It is vital that the UK makes it clear to those who develop and deploy cyber weapons that they must ensure compliance with IHL. The UK may even consider creating obstacles to make repurposing difficult and expensive. While it is hardly possible from a technical standpoint to guarantee that malware cannot be repurposed, methods like encrypting its payload and including obstacles in different components of the code, for example, could raise the bar in terms of the expertise required to re-engineer malicious tools that could have indiscriminate and harmful effects.

9. In understanding the harm caused by attacks within armed conflict, and in assessing whether incidental civilian harm is excessive to the military objective sought, it is the ICRC's position that the foreseeable indirect, or reverberating effects of an attack must also be considered in this assessment. This would be the case in cyber, as it would be in more conventional warfare. For example, in relation to a cyber operation on an electricity network that results in cutting off a hospital's electricity supply, the harm would include the foreseeable death of patients in intensive-care units of that hospital. Overall ensuring that civilian casualty monitoring and mitigation processes are adapted to the particular characteristics of cyberspace is crucial to ensure civilian harm is kept to a minimum.

10. Cyber operations are one of the areas where states explore the so-called grey zone to a maximum. In the ICRC's view, there is no doubt that IHL regulates the use of cyber operations during armed conflicts. On the other hand, most cyber operations are carried out below the threshold of armed conflict; to call this war risks undermining the more restrictive rules of international law applicable in peacetime, in particular international human rights law.

11. 2021 presents an opportunity for the UK Government to establish itself as a leading diplomatic power in setting standards, not only via the path set out in the IR but through its leadership of the G7. Substantial work has already been done in this forum, and we look forward to the UK's further contributions on this matter.

b) <u>Cyber Security</u>

12. As with offensive cyber, we noted the commitment to cyber security in the IR. As part of our long-standing, bilateral dialogue with the UK to reduce the human cost of armed conflict and other situations of violence, we look forward to contributing to the subsequent National Cyber Security Strategy in the summer. The IR makes clear that the UK is well placed and resourced to set standards. As cyber space will only increase as a home for so many parts of industry and government, we recommend cyber governance focuses on protection of civilian infrastructure, given the aforementioned potentially catastrophic consequences of its failure.

13. The ICRC is deeply concerned with the potential human cost of cyber-attacks against critical infrastructure that provide essential services to populations, in particular the health-care sector. During the current pandemic, cyber-attacks, have affected the medical sector in several states.[5] These attacks have come at a time when medical facilities and staff are

[5] https://www.icrc.org/en/document/governments-work-together-stop-cyber-attacks-health-care

under immense pressure. Cyber operations that disrupt hospital computers, medical supply chains, or medical devices pose great risk to those seeking medical care. Another disturbing target has emerged in vaccine programmes, on which we are all reliant to exit the acute phase of the COVID-19 pandemic. This is despite the protection afforded by international law to vaccine development programmes

14. In the ICRC's view, the real threat of cyber-attacks as seen during this pandemic must be a wake-up call. The health-care sector is among the most essential services in every state – not only in times of a health crisis, but at all times. This example demonstrates where leadership on encouraging responsible use of technology has a 'real world' effect on a crisis that has impacted us all.

15. The humanitarian imperative behind robust cyber security means we welcome the Foreign Secretary's commitment of the UK to cyber capacity building in growing economies, which are not only vulnerable to cyber-attacks, but may suffer worse consequences. We also noted with interest the Capacity Building Capability[6]. This commitment brings responsibility, particularly where offering this to parties to conflict. The UK should ensure all cyber security capacity building is conflict sensitive and focused on minimising harm to civilians. Building national-level cyber security is not an end in itself. Governments should be supported to create a resilient cyber ecosystem, which protects people and enables economies. Strong legal and normative frameworks that include IHL are essential for protection from predatory domestic authorities, foreign interference and cyber threats.

Data/Digital/AI/Machine Learning – The Human Centric approach

"Digital technologies and artificial intelligence are transforming the way people and organisations function in both the physical and virtual worlds. Digitalisation is also altering the way states, non-state armed groups (NSAGs) and other actors protect or restrict fundamental rights, and also how they manage security and conduct warfare. While we may not be able to predict where technological progress will lead us, we know that we must equip ourselves to understand its exponentially increasing impact on our environment, so that we can exploit the opportunities it offers and mitigate the risks it carries." (ICRC Strategy, 2019-22)

a) Data Protection

16. To many, data remains an abstract concept. For each individual the amount of data attributed to us, our families, our workplaces is only growing – as is the reliance on data used by both industry and Government. The ICRC has worked with sensitive data since our inception, e.g. holding as we do files on many, many missing people in conflicts today, and from those of decades ago. As such, we are keenly aware of the fact that risk exists in moving data onto digital platforms, and have developed respect for the potential consequences, along with ongoing attempts to mitigate them. As with cyber security, there are specific risks to populations affected by conflict, and also specifically severe consequences. As our data protection policy states: "The ICRC is concerned that unsuitable usage of new technologies could lead to stigmatisation, increased vulnerability and fragility, discrimination, persecution, and attacks on the physical and psychological integrity of

[6]https://www.gov.uk/government/speeches/cyberuk-conference-2021-foreign-secretarys-speech

certain populations in insecure environments. In this sense, the proper use of new technologies is a matter of life and death."[7]

17. The digital age has brought with it an eruption of innovations, many of which have rapidly become globally transformative. It is inevitable and admirable that governments, the technology industry and the humanitarian sector have sought to introduce these technologies to alleviate the suffering of some of the world's most vulnerable people. However, new technologies such as blockchain are embraced as a panacea to the problems of protracted conflicts, hard-to-reach populations or prevention of fraud. While we do not question their usefulness, where applied in the humanitarian context, with the risks being almost unknowingly high, these technologies create a greater necessity for caution as regards the collection, storage and use of data involved.

18. A key concern, one which arises from the complexity of the technologies themselves, involves an ever greater number of stakeholders, and margins for error, across a greater number of jurisdictions. The agreement between a humanitarian organisation and an individual who requires our help is, like the social contract between government and citizen, fundamentally one of trust. As such, we remain deeply concerned by technology that can take sensitive data potentially beyond our control.

19. The UK Government bears similar responsibility to vulnerable people across the world given the risks they carry. A cross-sectoral effort is needed to understand risks and urgent attention is required to conceptualise, track and mitigate harms. We recommend that all those with technology capable of collecting, storing and utilising data do not use new technologies until they have demonstrated that they understand the risks involved in them and have implemented mitigating measures.

20. The ICRC recently contributed to the International Development Committee's inquiry on racism within the Aid Sector and noted the particular risk in using algorithms reliant on personal data. The reliability of the results of such processes is dependent on the data put in. In short, biased data will lead to biased results.[8]

21. Again, with the UK Government seeking to use these new technologies, we recommend advocacy for identifying and tackling biases in the UK's Artificial Intelligence community. Similar initiatives are being rolled out in the humanitarian sector and in affected populations[9] but a cross-sectoral approach is necessary to truly mitigate this risk.

b) <u>Artificial Intelligence (AI) and Surveillance</u>

22. AI has the potential to be truly helpful 'on the ground' with affected populations. A key example of this is in cash programming, now almost exclusively delivered on digital platforms, in which those impacted by conflict are given the means with which to support

[7] https://www.icrc.org/en/document/icrc-data-protection-framework
[8] https://committees.parliament.uk/writtenevidence/26851/pdf/
[9] https://international-review.icrc.org/articles/ai-humanitarian-action-human-rights-ethics-913

themselves directly.[10] Much like the benefits of the M-Pesa application that the Foreign Secretary highlighted in the UKCYBER speech, cash programming has empowered people and protected dignity. From a purely humanitarian perspective, it has mitigated the negative secondary effects of in-kind assistance and boosted operational efficiency.[11]

23. However, there is a potentially serious adverse side to the use of these technologies and that is where something of assistance to the user becomes a means of surveillance of the user. Again, where such platforms can (and are) often used for profit, there needs to be real understanding of how the user is being instrumentalised to turn this profit. Digital payments are linked to identities and personal information; this information can identify, exclude and discriminate.

24. Neutrality is central to principled humanitarian action[12] and digital platforms are not neutral, being aligned with the political and financial objectives of those who created them. While they can be used for objectively good outcomes, we refer again to trust: principled humanitarian action is rooted in trust and humanitarians have a responsibility to gain the trust of the affected populations they aim to serve. There is inevitably then a tension between the neutrality of the humanitarian actor and the business model of the platform. A baseline for managing this risk is ethical and responsible use of people's data.

25. As the UK Government positions itself as a technological leader in using these innovations to bring benefit to some of the world's most disadvantaged, we strongly advocate that an ethical framework is applied which does not a) do harm to the user of the platform and b) breach the trust between those receiving assistance and those giving it. Again, in conflict-affected populations, there can be specific, severe humanitarian consequences if trust is undermined by technology. This must be assessed through a conflict-specific lens in policy-making.

c) <u>Misinformation, Disinformation & Hate Speech (MDH)</u>

"*Digital platforms such as Facebook, Alipay and WhatsApp may have more users ('virtual residents') than the populations of most countries; they operate quasi-global infrastructures, act as cross-border 'content policemen' and have market capitalizations that dwarf other sectors and most national GDPs."*[13]

26. While the ICRC is used to working with affected populations we, like many others, are adapting to working with the "virtual residents" forming the online populations articulated above. These virtual populations, if impacted by conflict and violenc,e are especially vulnerable to the misuse and mismanagement of digital platforms to spread MDH.

[10] https://international-review.icrc.org/articles/doing-no-harm-digitalization-of-cash-humanitarian-action-913
[11] https://blogs.icrc.org/law-and-policy/2021/03/02/cashless-cash/
[12] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/659965/UK-Humanitarian-Reform-Policy1.pdf
[13] "The changing role of multilateral forums in regulating armed conflict in the digital age" Amandeep S. Gill.

27. The rapid evolution of digital information technologies is turning MDH into an exacerbating and accelerating driver of conflict dynamics, violence, and harm – and an area of growing humanitarian concern.

28. There are direct and indirect humanitarian consequences to the spread of MDH, e.g. incitement to violence against a minority group, or via online harassment causing psychological harm – these can result in persecution, discrimination or displacement.[14]

29. People in situations of war are especially vulnerable to MDH. Discussions around this subject tend to focus on the impact on democracy. While incidents such as the protests on Capitol Hill and the pandemic have brought greater attention to the issue, governments now have an opportunity and responsibility to close this gap and prioritise conflict-sensitivity across policy making.

30. Social media can be a powerful positive force, but a cross-sectoral approach is needed for true mitigation of the risks and harms caused by the spread of online MDH. The power outlined above that sits within these "virtual populations" and their platforms means the UK will need to work closely with allies and big technology companies.

31. The ICRC has been encouraged by the UK's commitment to the G7 to address MDH. Again, this needs to also be looked at with distinct conflict sensitivity and we encourage ongoing work with humanitarian organisations to achieve this. While we will soon be sharing our report on MDH in conflict and situations of violence much more work needs to be done – with the support of Governments – to understand the humanitarian impacts.

d) <u>Autonomous Weapon Systems (AWS)</u>

32. As of May 12 2021, the ICRC updated its position on autonomous weapon systems (AWS). The ICRC understands an AWS to be a weapon system that selects and applies force to targets without human intervention, in the sense that after initial activation by a person, an AWS self-initiates or triggers a strike in response to information from the environment received through sensors and on the basis of a generalized 'target profile'. This understanding of an AWS is broader than the definition proposed by the UK MoD in 2017.[15]

33. In the process by which AWS function – the user does not choose, or even know, the specific target(s) and the precise timing and/or location of the resulting application(s) of force. This:

- brings risks of harm for those affected by armed conflict, both civilians and combatants, as well as dangers of conflict escalation

- raises challenges for compliance with international law, including IHL, notably the rules on the conduct of hostilities for the protection of civilians

- raises fundamental ethical concerns for humanity, in effect substituting human decisions about life-and-death with sensor, software and machine processes.

34. In light of these risks, and current trends in military interest and investments that are set to accentuate them, the ICRC has, since 2015, urged states to establish internationally

[14] https://blogs.icrc.org/law-and-policy/2021/03/30/fog-of-war-and-information/
[15] https://lordslibrary.parliament.uk/killer-robots-should-lethal-autonomous-weapons-be-banned/

agreed limits on AWS. With a view to supporting current efforts to establish such limits, the ICRC now recommends that states adopt new legally binding rules, including, in particular,

- a prohibition on 'unpredictable AWS', that is, AWS that are designed or used in a manner such that their effects cannot be sufficiently understood, predicted and explained

- a prohibition on 'anti-personnel AWS', that is, AWS that are designed or used to apply force against persons

- regulations on the design and use of all other AWS, including a combination of limits on the types of targets; limits on the duration, geographical scope and scale of use; limits on situations of use; and requirements for human-machine interaction.

35. There are significant areas of overlap between the ICRC's and the UK Government's positions on AWS, including shared recognition of the importance of appropriate human judgment and control, and the consequent need to set constraints on the design and use of AWS. It is an open question to what extent the ICRC's recommended prohibitions overlap with the UK Government's notion of 'fully AWS' whose development the Government rules out. Whereas the UK Government is not supporting the adoption of new legal rules on AWS to date, deeming existing IHL sufficient to address concerns raised by AWS, it sees value in compiling good practice guidance. Beyond new legal rules on AWS, the ICRC supports initiatives that aim at effectively addressing concerns raised by AWS in a timely manner.

Conclusion

36. Even where new technology is applied through one specific context, such has been done in this paper, its cross-cutting nature and the extent of the risks it poses is very clear. New technology is shaping conflict and having clear humanitarian consequences in exceptionally vulnerable populations. As can be seen above, artificial intelligence runs through these concerns as a single and powerful narrative which demands responsible, ethical and cautious use. We are strongly encouraged by the UK Government's commitment to good global governance in respect of the cyber space and wider technology. We continue to emphasise the need to also develop a conflict-sensitive approach, to ensure this technology is used for the benefit of humanity, rather than to its detriment. The COVID-19 crisis has provided a stark example of where science and technology can truly be used for global good, and we look forward to working further with the UK Government to ensure that humanitarian concern is at the core of their policy around new technology.

June 2021