



红十字国际委员会 关于自主武器系统的立场

红十字国际委员会关于自主武器系统的立场.....	2
背景文件.....	3
1. 国际上关于自主武器系统的探讨.....	3
2. 当前及新兴的自主武器系统.....	5
3. 需要对自主武器系统施加的限制.....	7
3.1 应对自主武器系统的不可预测性所引发的关切.....	7
人道关切.....	7
国际人道法关切.....	7
红十字国际委员会建议：禁止使用不可预测的自主武器系统.....	7
3.2 应对使用自主武器系统攻击人类所引发的关切.....	8
道德关切.....	8
国际人道法关切.....	9
红十字国际委员会建议：禁止使用杀伤人员自主武器系统.....	9
3.3 应对其他自主武器系统所引发的关切.....	9
人道、法律和道德关切.....	9
当前实践中用于降低风险的措施类型.....	10
红十字国际委员会建议：规制其他自主武器系统.....	10
4. 结论及红十字国际委员会对各国建议的总结.....	11

红十字国际委员会关于自主武器系统的立场

红十字国际委员会关于自主武器系统的关切

自主武器系统在没有人类干预的情况下选择目标并对其使用武力。在最初由人启动或发射后，自主武器系统根据从传感器接收到的环境信息，并基于概括性的“目标描述”，自行启动或触发攻击。这意味着使用者并未选择自主武器系统使用武力攻打的具体目标以及精确的时间和(或)地点，甚或对此一无所知。

使用自主武器系统会带来风险，是因为其影响难以预测，且难以限制，人类由此丧失了对武力和武器使用的控制和判断。这引发了人道、法律和道德方面的严重关切。

自主武器系统的运作过程：

- **可能会为受武装冲突影响之人（无论是平民还是战斗员）带来伤害**，而且具有导致冲突升级的危险；
- **对遵守国际人道法等国际法规则**，特别是保护平民的敌对行动规则，**提出了挑战**；
- 实际上用传感器、软件和机器程序取代了人类对生死的决定，因而引发**人类的根本道德关切**。

红十字国际委员会就规制自主武器系统向各国提出的建议

自2015年以来，红十字国际委员会一直敦促各国就自主武器系统确立国际公认的限制标准，以确保对平民的保护、对国际人道法的遵守，以及道德上的可接受性。

为支持当前工作，就自主武器系统确立国际限制标准，从而应对其带来的风险，**红十字国际委员会建议各国通过具有法律约束力的新规则**。特别是：

- 1 对不可预测的自主武器系统应予以明令禁止**。提出这一规则，一项突出原因就是此类系统会造成不分皂白的后果。要实现这一点，就最好禁止因设计或使用方式而使人无法充分理解、预测并解释所造成影响的自主武器系统。
- 2 鉴于维护人道，以及捍卫国际人道法中保护平民和失去战斗力之战斗员的规则的道德考量，应禁止使用自主武器系统攻击人类**。要实现这一点，就最好禁止旨在或用于对人类使用武力的自主武器系统。
- 3 为了保护平民和民用物体、捍卫国际人道法规则并维护人道，对于未被禁止的自主武器系统，其设计和使用时应受到规制**，包括通过综合以下各项限制予以规范：
 - **对目标类型的限制**，例如仅限于攻打属于军事目标性质的物体；
 - **对使用时长、地理范围和规模的限制**，包括允许对特定攻击进行人为判断和控制；
 - **对使用情形的限制**，例如仅限于在不涉及平民或民用物体的情况下使用；
 - **对人机互动的要求**，特别是要确保有效的人类监督，以及及时的干预和停用。

为有效解决自主武器系统引发的关切，各国提出了多项倡议，旨在确立国际标准以对其施以限制，例如各国在《特定常规武器公约》框架下为商定规范和行动框架的多方面问题开展了相关工作。红十字国际委员会对这些倡议表示支持。考虑到自主武器系统的技术和使用发展迅速，及时确立国际公认的限制标准至关重要。此类标准除了新的法规，还可包括共同政策标准和良好实践指南，二者相互补充、相辅相成。为此，红十字国际委员会在其职责和专业范围内，随时准备与政府、武装部队以及科技界和产业界代表等国际和国家层面的利益相关方开展合作。

2021年5月12日，日内瓦

背景文件

1. 国际上关于自主武器系统的探讨

过去10年间，国际上一直在探讨自主武器系统所引发的人道、法律和道德关切。《特定常规武器公约》缔约国从2014年起就开始讨论自主武器系统问题，后又自2016年开始通过正式成立的致命性自主武器系统领域的新兴技术问题政府专家组开展工作。

2019年，《特定常规武器公约》缔约国同意就自主武器系统“规范和行动框架的多方面问题”达成共识建议，同时通过了11项指导原则，反映迄今为止所达成的共识。¹ 2020年，许多国家在提交给政府专家组的国家评论中，并在同年9月专家组会议审议期间，阐述了对这些原则的理解。正如此次会议期间及会后历任政府专家组主席所指出的，这表明各国之间的意见日益趋同。² 2021年，政府专家组将进一步组织会议，迎接《特定常规武器公约》第六次审议大会——这是缔约国应对自主武器系统所引发关切的关键时刻。

红十字国际委员会于2011年首次公开就自主武器系统问题表示关切，引起各界注意。**2015年以来，委员会一直呼吁各国迫切需要针对自主武器系统确立国际公认的限制标准**，以应对自主武器系统使用方面的迅速发展扩张，及其引发的人道、法律和道德关切。此后，委员会向各国就自主武器系统一般所需的限制类型提出建议，尤其是针对此类武器的可预测性、攻击目标类型、使用时长和范围、使用情形、以及人类监督方面的限制。最近，委员会还在对《特定常规武器公约》政府专家组指导原则的评注中提出了建议。³ 截至目前，关于上述限制标准是应当通过制定具有法律约束力的新规、颁布政策标准还是采纳共同实践等形式予以体现，委员会尚未给予解答。

红十字国际委员会的立场及其向各国提出的建议基于以下三个方面：一是委员会对自主武器系统相关的人道、法律、道德、技术和军事影响的分析；二是委员会发表在一系列报告中的见解，如2020年6月与斯德哥尔摩国际和平研究所(SIPRI)联合发表的题为《对武器系统自主性的限制：明确人类控制的实践要素》的报告；三是与各国在《特定常规武器公约》专家会议上的定期接触探讨以及与各国展开的双边合作。⁴

1 联合国，《〈禁止或限制使用某些可被认为具有过分伤害力或滥杀滥伤作用的常规武器公约〉缔约方会议》，日内瓦，2019年11月13日至15日，最后报告，联合国CCW/MSP/2019/9号文件，2019年12月13日。

2 联合国，《致命性自主武器系统领域的新兴技术问题政府专家组：各国对指导原则之评论的共同点》，联合国CCW/GGE.1/2020/WP.1号文件，2020年10月26日；联合国，《致命性自主武器系统领域的新兴技术问题政府专家组：主席的总结》，联合国CCW/GGE.1/2020/SP.7(预备稿)，2021年4月19日。

3 ICRC, *ICRC Commentary on the “Guiding Principles” of the CCW GGE on “Lethal Autonomous Weapons Systems”*, July 2020.

4 ICRC, *Statement of the ICRC to the UN CCW GGE on Lethal Autonomous Weapons Systems*, 21–25 September 2020, Geneva; ICRC, *ICRC Commentary on the “Guiding Principles” of the CCW GGE on “Lethal Autonomous Weapons Systems”*, July 2020; V. Boulanin, N. Davison, N. Goussac, and M. Peldán Carlsson, *Limits on Autonomy in Weapon Systems: Identifying Practical Elements of Human Control*, ICRC & SIPRI, June 2020; 红十字国际委员会，《国际人道法及其在当代武装冲突中面临的挑战》，第33届红十字与红新月国际大会，日内瓦，2019年10月，第29~31页；ICRC, *Autonomy, Artificial Intelligence and Robotics: Technical Aspects of Human Control*, August 2019; ICRC, *Statements of the ICRC to the UN CCW GGE on Lethal Autonomous Weapons Systems*, 25–29 March 2019, Geneva; ICRC, *The Element of Human Control*, working paper submitted at the Meeting of High Contracting Parties to the CCW, Geneva, 21–23 November 2018, CCW/MSP/2018/WP.3, 20 November 2018; ICRC, *Ethics and Autonomous Weapon Systems: An Ethical Basis for Human Control?*, 3 April 2018; 红十字国际委员会，《红十字国际委员会关于自主武器系统的观点》，2016年4月11日；红十字国际委员会，《自主武器系统：增强武器关键功能的自主性带来的影响》，2016年3月；红十字国际委员会，《自主武器系统——技术、军事、法律和人造视角》，2014年3月。

在此基础上，红十字国际委员会如今可以**提供更详细的建议，说明**为确保对平民的保护、对国际人道法的遵守，以及道德上的可接受性，**需要对自主武器系统施加哪些具体限制**。此外，**委员会深信，要实施上述限制标准，应具体针对自主武器系统制定具有法律约束力的新规则，从而加以规范**。这些规则应澄清国际人道法等现有国际法规则如何限制自主武器系统的设计和使用，并在必要时补充法律框架，包括应对自主武器系统所引发的更广泛的人道风险和根本道德关切。

就自主武器系统具有法律约束力的新规进行谈判，以及为制定《特定常规武器公约》政府专家组所考虑的行动和规范框架的各个方面而开展其他工作，⁵可以是相辅相成的。相关工作的一部分就是旨在有效应对自主武器系统所引发关切的行动。为实现这一目标，各国可以在政治宣言中就国际承诺达成一致，其他手段还包括对测试、验证或核查的国际技术标准详加阐述、由国家自主武器系统的开发或采购进行延期偿付，以及采取措施，支持在国内，包括军事条令和其他指南中实施国际公认的限制标准。

⁵ 联合国，《〈禁止或限制使用某些可被认为具有过分伤害力或滥杀滥伤作用的常规武器公约〉缔约方会议》，日内瓦，2019年11月13日至15日，最后报告，联合国CCW/MSP/2019/9号文件，2019年12月13日。

2. 当前及新兴的自主武器系统

红十字国际委员会认为，自主武器系统是在没有人类干预的情况下选择目标，并对其使用武力的武器。在最初由人启动或发射后，自主武器系统根据从传感器接收到的环境信息，并基于概括性的“目标描述”，自行启动或触发攻击（泛化的技术指标代替了具体的目标）。

简而言之，自主武器系统在被物体或人触发时会自行发动攻击，但使用者并不确切知晓此类武器会在何时攻击何处，也并未就此做出选择。事实上，要理解非自主武器系统和自主武器系统之间的区别，就要看武器所攻击的特定目标是否由人来选择。⁶有很多武器系统、平台和弹药，尤其是目前通过遥控操作的无人系统，都能实现自主使用武力的功能。

一些自主武器系统已经在狭义情况下用于特定任务，例如：防空系统在军舰上或军事基地用于攻击来袭导弹、火箭弹或迫击炮；坦克上的“主动防御”武器用于打击同类来袭炮弹；具备自主模式的巡飞弹用于攻击雷达（可能还有车辆）；以及某些导弹和传感器引爆弹药用于攻打军舰和坦克。地雷也被称为粗制的自主武器系统。⁷在支持者看来，与直接控制和遥控武器系统相比，自主武器系统具备以下若干潜在的军事利益，包括：

- 攻击速度提升：加快对目标的探测、跟踪和使用武力的过程。这提供了军事利益，但有可能导致对武力使用丧失控制，造成战斗升级。
- 自动区域拒止：自主武器系统可以阻止敌方进入或穿过某些区域，而无需派驻士兵或进行持续监视。这与布置雷区的军事原理相似。
- 在通讯中断时可继续攻击：遥控（海陆空）武装无人机的操作人员如要发起攻击，需要依赖通信连接，但通信线路容易受到干扰、切断或遭遇黑客攻击；自主武器系统则可以在没有通讯的情况下运作。
- 大量、集群作战：在进行单次攻击时，由于使用自主武器系统，无需操作人员参与，这就有助于在比遥控系统所需人力更少的情况下部署更多的无人武装系统。

一些支持者还声称，他们正力图研制在攻击目标方面比直接控制武器或遥控武器（非自主武器系统）精度和（或）准确度更高的自主武器系统。但实际上，自主武器系统由于在攻击目标上转变为更为泛化的决策方法，对最终目标以及使用武力的确切时间和（或）位置了解较少，攻击的精度和准确度也因此遭到削弱。不过，限制使用自主武器系统，并不代表军方不能使用新技术来确保提升目标选择的精度和准确度。

支持者提出的另一个常见论点是，在遵守国际人道法方面，使用自主武器系统将“优于人类”。然而，要评估自主武器系统所造成的风险，我们不需要将其与人类比较；相反，我们需要比较的是（a）人类使用非自主武器系统攻击人类所选择目标的后果，以及（b）人类使用自主武器系统攻击并非人类专门选择的目标的后果。无论人类决策者如今在根据国际人道法预测并限制自身攻击所造成的影响方面面临何种挑战，鉴于自主武器系统的运作方式，这些挑战都因此加剧，而并非减少。

就使用自主武器系统而言，现有的军事实践侧重于施以严格限制，有助于使平民和“友军”免遭风险，促进遵守国际人道法，而且此类限制还有可能受到道德考量的影响。限制对象包括：

- 攻击目标：自主武器系统通常用于攻击军用物体，如炮弹、飞机、海军舰艇、军用雷达、坦克或其他军用车辆。据红十字国际委员会所知，目前杀伤人员自主武器系统并未投入使用（但为《禁止杀伤人员地雷公约》所禁止，且受到《特定常规武器公约经修正之第二号议定书》规制的杀伤人员地雷除外）。

6 红十字国际委员会，《国际人道法及其在当代武装冲突中面临的挑战》，第33届红十字与红新月国际大会，2019年10月，第29~31页。

7 V. Boulanin, N. Davison, N. Goussac, and M. Peldán Carlsson, *Limits on Autonomy in Weapon Systems: Identifying Practical Elements of Human Control*, ICRC & SIPRI, June 2020, p. 18. 另见红十字国际委员会，《自主武器系统：增强武器关键功能的自主性带来的影响》，2016年3月，第7页。

- 使用时长和地理范围：大多数自主武器系统仅在短时间内处于自主模式，并且很多武器不可移动，而是固定于一处。
- 使用情形：大多数自主武器系统仅在不涉及平民和民用物体的情况下，或在采取措施（如设置障碍、警示标志、禁区），阻止平民进入自主武器系统运作区域的情况下使用。
- 人机互动：几乎所有自主武器系统都由人类操作员实时监控，并在必要时加以干预，包括对其运作进行授权，推翻、否决其决定或将其关闭。

然而，武器系统的基础设施正在不断扩张，其中有大量系统可能会成为自主武器系统，从带有面部识别功能的手持式武装四轴飞行器到自主战斗机，从“哨兵枪”到自主坦克，再从武装快艇到自主猎潜水下无人机，应有尽有。基础设施中设有相互连接的网络系统，其中用于识别和选择目标的软件可分别触发不同的武器，以及自主网络武器。

许多遥控系统已经可以自主识别、跟踪或选择目标，而要想使其自主使用武力，只需投入低廉的成本进行软件升级，甚或只需改变条令。武器如果发生故障或遭到黑客蓄意攻击，也有可能发生这种情况。例如，在某些边境地区和军事基地部署的遥控“哨兵枪”就被用于自主选择人类目标。据红十字国际委员会所知，使用者仍然必须通过遥控的方式专门对使用武力进行授权，不过在商业层面，已经开发出了自主式哨兵枪。

当前军事利益和投资的趋势表明，如果不确立国际公认的限制标准，未来的自主武器系统可能会：

- 日益依赖人工智能和机器学习软件，由此令人担忧其在设计层面就存在不可预测性；
- 用于攻击人类，且攻击的物体种类增多；
- 在移动性方面不断增强，攻击范围更广，使用时间更长，并能发起多次攻击；
- 用于城镇地区，致使平民陷入最为危险的境地；
- 在使用时缺乏有效的人类监督和及时的干预或停用。

上述趋势不仅限于资源充足的国家，而且还是当前军事技术和条令迅速发展，以及自主武器在国家和非国家武装团体中不断扩散所体现出的特点，使得下一节所概述的人道、法律和道德关切愈发深重。考虑到上述趋势，更是迫切需要就针对自主武器系统具有法律约束力的新规，以及《特定常规武器公约》政府专家组审议中的自主武器系统规范和运作框架的其他方面，在国际上达成一致意见。

3. 需要对自主武器系统施加的限制

鉴于自主武器系统的运作过程，人类对武力和武器的使用丧失了控制和判断，这在人道、法律和道德层面均引发严重关切。一般而言，使用自主武器系统会有损对平民的保护、挑战法治，并就人道原则引发关切，从而大大增加受武装冲突影响民众所面临的风险。

自主武器系统作为一种作战手段，必须能够使用，且必须按照国际人道法的规定加以使用。而要遵守国际人道法关于敌对行动之规定的，必须是自主武器系统的使用者，而非武器本身。负责适用国际人道法，且能对违反国际人道法的行为负责的，也是武装冲突各方，归根结底，即为人类。⁸然而，**自主武器系统运作的过程对遵守相关国际人道法规则提出了挑战。**

3.1 应对自主武器系统的不可预测性所引发的关切

人道关切

由于自主武器系统具体在何时(或)何地攻击何种具体目标并非由使用者所选择，也并非为使用者所知晓，**因此，使用所有自主武器系统所造成的影响都必然在一定程度上是不可预测的。**这使受武装冲突影响的民众有可能遭到伤害，对适用国际人道法构成严重挑战，并有可能导致冲突升级。

本文第2节中所提出的趋势(具体而言，即自主武器系统打击目标的种类更多；使用时间更长，攻击范围更广；使用环境更加多变、拥挤和复杂；且人类参与减少)**将使得自主武器系统的影响更加难以预测，从而增加平民的风险。**

此外，**开发由人工智能，尤其是机器学习软件控制的自主武器系统，更是在设计层面额外增加了不可预测性。**无论在何种环境下使用，由于机器学习技术，人类都很难理解自主武器系统的运作过程，因而也无法做出预测和解释(“黑箱”挑战)。⁹

国际人道法关切

自主武器系统的不可预测性对国际人道法提出了根本挑战。习惯国际人道法禁止使用不分皂白性质的武器，即在正常或预期使用情况下不能以特定军事目标为对象，或其效果不能按照国际人道法的要求加以限制的武器。¹⁰

某些自主武器系统本身就属于不分皂白的性质，并因此被现有国际人道法所禁止。其中尤其要指出的一类，就是在正常或预期使用情况下无法对其影响予以充分理解、预测和解释的自主武器系统。例如，如果对使用自主武器系统负责的人员无法合理预期何种因素会触发自主武器系统进行攻击，他们就无法按照国际人道法的要求控制和限制其影响，也无法解释为何某人或某物体因自主武器系统所遭受的攻击会使他们自身违反国际人道法，并需对违法行为负责。

具体而言，如果自主武器系统的运作并非透明公开，则负责适用国际人道法规则的人员，亦即受委托对自主武器系统进行法律审查的人员，和负责在此类系统使用期间遵守国际人道法的人员，就无法合理确定自主武器系统是否符合国际人道法。运作过程不透明，主要是由于这一过程依赖于人工智能和机器学习技术，或是由于其在使用过程中发生了变化，并对武力使用产生了影响(例如，借助机器学习，攻击行为的参数能够随时间推移发生变化)。

红十字国际委员会建议：禁止使用不可预测的自主武器系统

根据上述分析，**对不可预测的自主武器系统应予以明令禁止。提出这一规则，一项突出原因就是此类系统会造成不分皂白的后果：**使用者无法知晓此类武器是会攻击平民还是战斗员，民用物体还是军用

8 红十字国际委员会，《国际人道法及其在当代武装冲突中面临的挑战》，第33届红十字与红新月国际大会，日内瓦，2019年10月，第29-31页。

9 ICRC, *Autonomy, Artificial Intelligence and Robotics: Technical Aspects of Human Control*, August 2019; 红十字国际委员会，《武装冲突中的人工智能与机器学习：以人为本的方法》，2019年6月。

10 红十字国际委员会，《习惯国际人道法第一卷：规则》，规则71，2005年。

物体，也无法得知其影响是否会按照国际人道法的要求受到限制。**要明确这一点，最好是禁止由于设计或使用方式而使人无法对所造成影响予以充分理解、预测和解释的自主武器系统。**

这项禁令要得以施行，其基础是各国认识到，为遵守国际人道法，且出于实际的军事行动原因，需要能够对自主武器系统的使用做出充分的预测。而其支撑依据在于，现有国际人道法禁止使用本质上不分皂白的武器，这是普遍达成一致的观点。另外，对不可预测的自主武器系统通过条约施以禁令，也会有助于明确哪些自主武器系统会被视作不分皂白的武器。

3.2 应对使用自主武器系统攻击人类所引发的关切

正如红十字国际委员会¹¹和其他相关方先前所强调的，旨在或用于攻击人类的自主武器系统还会引发特殊的道德关切和法律挑战。

道德关切

自主武器系统的运作过程实际上是用传感器、软件和机器程序取代了人类对生死的决定，因而**引发人类的根本道德关切。总而言之，大多数人都认为算法，也就是机器程序，不应决定人的生死。**不过，出于这一关切，到底是应当禁止使用所有自主武器系统，禁止危害人类的自主武器系统，还是只禁止直接攻击人类的自主武器系统，目前并不总是有确切的答案。

许多国家、¹²联合国秘书长、¹³民间社会¹⁴以及科技界的主要人物¹⁵都提出了上述关切。

这些关切重点关注的是在生死抉择上，人类能动性、道德责任和人的尊严相互关联，一失俱失。就决策和行动而言，人类有道德机制和责任作为指引，但无生命物体（如武器、机器和软件）却没有。无论自主武器系统如何“精密”，情况也依然如此。

要维护人类能动性，就需要人类有效的审慎思考。否则，就可以说一直以来所做的决策均未负起道德责任，受到攻击或影响之人的人格尊严也从未得到承认。而去除人类能动性，就等同于非人化，损害了共同的人性，在关乎生死的决定中还会使人无法实施克制。而人性中的克制就体现在，即使属于合法情况，人也有可能决定不使用武力。

红十字国际委员会认为，上述道德关切适用于危害人类的自主武器系统，且就旨在或用于直接攻击人类的自主武器系统（相对于攻击导弹等无人操纵的军用物体的自主武器系统）而言最为深重。这是因为后者漠视了人类宝贵的生命，将其视同传感器数据和一系列机器程序，助长了仅基于概括性的目标描述就加以攻击，并致死致伤的行为。¹⁶这实际上相当于“算法致死”。如不坚守住这一最后阵地，人类下一步就将迎来自动杀戮。

¹¹ ICRC, *Ethics and Autonomous Weapon Systems: An Ethical Basis for Human Control?*, 3 April 2018.

¹² See, V. Boulanin, N. Davison, N. Goussac, and M. Peldán Carlsson, *Limits on Autonomy in Weapon Systems: Identifying Practical Elements of Human Control*, ICRC & SIPRI, June 2020, p. 10, note 22.

¹³ UN Secretary-General, “Machines Capable of Taking Lives without Human Involvement Are Unacceptable, Secretary-General Tells Experts on Autonomous Weapons Systems”, SG/SM/19512-DC/3797, 25 March 2019.

¹⁴ e.g. Human Rights Watch, *Losing Humanity: The Case against Killer Robots*, 18 November 2012; Article 36, “Targeting People”, *Policy Note*, November 2019.

¹⁵ e.g. Future of Life Institute, *An Open Letter to the United Nations Convention on Certain Conventional Weapons*, 2017, and Future of Life Institute, *Autonomous Weapons: An Open Letter from AI & Robotics Researchers*, 2015 (26个国家的4502名人工智能与机器人研究人员、26215名其他科学家和专家，以及100家人工智能与机器人公司创始人及首席执行官签署了公开信，呼吁禁止使用自主武器系统，并制定相关法规); Google, *AI Principles*, 2018.

¹⁶ V. Boulanin, N. Davison, N. Goussac, and M. Peldán Carlsson, *Limits on Autonomy in Weapon Systems: Identifying Practical Elements of Human Control*, ICRC & SIPRI, June 2020, p. 14: “在自主武器系统用于攻击人类，以及平民面临附带风险的情况下，根本道德关切似乎确实得以凸显（不过军用飞机、车辆、建筑等涉及人员的军事目标也可能会引发此类关切。）”；ICRC, *Ethics and Autonomous Weapon Systems: An Ethical Basis for Human Control?*, 3 April 2018, p. 22: “在使用武力的决策中人类能动性丧失、道德责任淡化以及人类尊严丧失这三项道德关切相互关联，三者结合起来，会导致最为深远的后果，可能会阻碍杀伤人员自主武器系统的发展和使用，而且取决于破坏物体目标对人的生命所带来的风险，这甚至还会限制反器材系统的运用。”

国际人道法关切

从法律角度看，自主武器系统对受国际人道法保护之人构成了真正的风险，使其有可能遭到伤害。特别是，使用自主武器系统攻击人类会导致重大风险，即受保护的平民和丧失战斗力的战斗员可能会触发自主武器系统的攻击。

要有效保护丧失战斗力的战斗员/战斗人员，以及没有或不再直接参与敌对行动的平民，就需要由人类在具体攻击的情形下充分考虑相关局势、行为、意图及因果关系，做出艰难的法律评估。就如何根据国际人道法合法使用杀伤人员自主武器系统而言，有两项相互关联的挑战，使人难以对此进行设想。第一，平民可能参与敌对行动的方式极为多样，参与敌对行动的战斗员或平民在投降时或对受伤做出反应的方式也是如此；因此，要确定一个人是受保护人员，还是合法目标，是高度依赖于相关情境的，不适合以目标描述的形式做标准化处理。第二，上述法律特征可能会很快发生变化，亦即指挥官在发动攻击时对自主武器系统运作区域内人员的可攻击性做出的假设可能会在自主武器系统进行实际攻击前发生变化。与属于军事目标性质的物体相比，为使人免遭攻击而提供的法律保护更容易因实际情况而变化(见下文第3.3节)。

在如今的战斗局势中，越来越多的战事发生在多变且拥挤的城市地区，遵守区分原则和保护丧失战斗力之战斗员的规则本就面临巨大挑战，而引入自主武器系统攻击人类，只会加剧上述挑战。红十字国际委员会认为，在实际战斗情况中，很难设想使用自主武器系统攻击人类不会引发违反国际人道法的重大风险。

红十字国际委员会建议：禁止使用杀伤人员自主武器系统

考虑到保护人类的道德考量，以及维护国际人道法关于保护平民和丧失战斗力之战斗员的规则，**应禁止使用攻击人类的自主武器系统**。要实现这一点，最好是禁止旨在或用于对人类使用武力的自主武器系统。

这一禁令基于目前尚未有人使用自主武器系统直接攻击人类的实际情况，许多国家、科学家、哲学家、人权专家、民间社会和广大公众所表达的关切也为其提供了支撑。各方都认为人类决不能将生死交由机器决定。

要禁止由人触发的自主武器系统，可参考《禁止杀伤人员地雷公约》中对杀伤人员地雷的禁止性规定。这项针对杀伤人员自主武器系统建议实施的禁令将会树立非常重要的规范。

3.3 应对其他自主武器系统所引发的关切

使用任何自主武器系统，都必须遵守国际人道法旨在敌对行动期间保护平民和民用物体的规则，特别是区分原则、对不分皂白和不成比例攻击的禁止，以及在攻击中采取一切可能的预防措施的义务。即使在上述讨论以外的情况下，使用自主武器系统也会引发人道、法律和道德方面的关切，红十字国际委员会在此类相关情况中也建议禁止使用自主武器系统。

人道、法律和道德关切

使用自主武器系统的风险在于，使用者在发起攻击时所做的决定，包括对自主武器系统将予攻击的目标是否为军事目标，以及攻击的比例性如何等问题的判定，可能会由于情况的变化而失效。尤其是，如果攻击目标在法律上的军事目标属性可能会快速转变，那么在自主武器系统攻击时间加长、运作区域扩大、可发起攻击的次数增多，以及运作环境更加多变、拥挤或复杂的情况下，决定失效的风险还会进一步上升。虽然现有自主武器系统在设计和使用上通常是为了尽量减少这些风险，并促进遵守国际人道法，**但第2节中所提出的自主武器系统发展趋势都显示出在这些方面的风险可能日趋增加**。

而且由于相关趋势，自主武器系统使用者还更有可能面临风险，既无法认识到情况已经发生变化，需要暂停攻击，也无法及时干预，防止造成不利的人道后果，导致违反国际人道法的行为。

鉴于当代武装冲突在不断演变，且城市战日益增加，**自主武器系统的设计和使用如毫无限制，将造成重大的人道风险和违反国际人道法的风险**。

当前实践中用于降低风险的措施类型

人道、法律、道德和军事行动理论依据相互强化，严格限制当前实践中自主武器系统的设计与使用，并提供了针对自主武器系统所需限制类型的示例，从而使得人类能够对武力的使用进行充分控制和判断，并降低上述强调的风险。做到这一点，是通过将技术限制与条令限制相结合的方法：

- **自主武器系统所攻击的目标一般限于军事目标的法律定性相对稳定的物体，即性质上属于军事目标的物体**，如炮弹、军用雷达或海军舰艇。而在法律上判定其他物体是否为军事目标(例如，建筑物或车辆如被敌方用于军事行动，就可能成为军事目标¹⁷)通常高度取决于具体情况，因此，在自主武器系统运作区域内外观相似的物体(如外表完全一致，但分别由平民和军方使用的车辆)在性质上就可能存在差异，而且法律定性可能在自主武器发起攻击和实际进行攻击的较短间隔内就迅速发生转变(例如，截至自主武器系统发起攻击之时，某民用车辆一直由敌方用于军事行动，但在自主武器系统实际进行攻击时，敌方已停止将该车用于军事用途，就属于这种情况)。
- **自主武器系统的使用通常在空间、时间和武力规模上受到限制**，具体涉及自主武器系统可以使用武力的区域范围、运作时长以及可以进行攻击的规模或次数。这些限制标准旨在使自主武器系统使用者对周边局势具备必要的认识，从而预测攻击的影响，并能在发起攻击时合理确定此次攻击将能够合乎国际人道法的规定，另外还能降低攻击期间具体情况可能发生变化的风险，有助于在自主武器系统运作期间加以监督。
- **自主武器系统通常在不涉及平民和民用物体的地点使用**。就自主武器系统可以使用武力的区域而言，如若区域内平民和民用物体越多，则对平民造成伤害的风险就越高，原因有以下两点。首先，对于目标描述要求攻击军用吉普车或人员输送车的自主武器系统，汽车或公共汽车等民用物体可能就会将其触发。其次，如果平民和民用物体处于军事目标(如军用吉普车或人员输送车)内或其附近，也可能受到附带损害。
在不涉及平民和民用物体的地点，例如远离航道或捕鱼区的公海，或是可以有效合法阻止平民和民用物体进入的地区(例如隔离出军事基地，或设置禁飞区)，管理上述风险更为容易。相比之下，如果在城镇等多变、拥挤或复杂的平民环境中使用自主武器系统，会使平民极有可能遭到伤害。在这种环境下，国际人道法保护平民的规则能否得到遵守，令人尤为关切。同样在道德层面令人深感关切的是，在使用自主武器系统时，某些人员虽并非直接攻击的目标，但却由于机器程序或机器计算导致的意外或附带损害而丧命。
- **自主武器系统通常在人类的持续监督下使用，人类还可选择将其停用**。在自主武器系统的设计和使用方面所采取的措施(包括上述对目标、时间、空间、武力使用规模和使用情形的限制标准)旨在让使用者能够实时认识到周边局势，并保证在实践中，使用者能在必要时加以干预，并关闭自主武器系统。

这其中存在的风险是，本文第2节所提出的趋势，特别是在使用人工智能和机器学习来控制对目标的选择和武力的使用方面速度越来越快、规模越来越大、依赖程度越来越深，将削弱人类操作员理解所接收信息，审慎思考其选择，并根据人道、法律和道德原则及时采取行动的能力。这反过来又会降低使自主武器系统操作人员对所造成伤害和违反国际人道法行为负责的可能性。

红十字国际委员会建议：规制其他自主武器系统

根据上述分析，为了避免对平民和民用物体造成伤害，捍卫国际人道法规则并维护人道，**对于未被禁止的自主武器，其设计和使用应受到规制**，包括通过综合以下各项具有法律约束力的限制予以规范：

- **对目标类型的限制**，例如仅限于攻打属于军事目标性质的物体；
- **对使用时长、地理范围和规模的限制**，包括允许对特定攻击进行人为判断和控制；
- **对使用情形的限制**，例如仅限于在不涉及平民或民用物体的情况下使用；
- **对人机互动的要求**，特别是要确保有效的人类监督，以及及时的干预和停用。

¹⁷ 日内瓦四公约《第一附加议定书》第52条；红十字国际委员会，《习惯国际人道法第一卷：规则》，规则7~10，2005年。

4. 结论及红十字国际委员会对各国建议的总结

红十字国际委员会认为，针对许多国家、民间社会和委员会强调的自主武器系统所引发的人道、法律和道德关切，迫切需要制定具有法律约束力的新规则。

为支持当前工作，就自主武器系统确立国际限制标准，从而应对其带来的风险，**红十字国际委员会建议各国通过具有法律约束力的新规则**。特别是：

- 1 对不可预测的自主武器系统应予以明令禁止。**提出这一规则，一项突出原因就是此类系统会造成不分皂白的后果。要实现这一点，就最好禁止因设计或使用方式而使人无法充分理解、预测并解释所造成影响的自主武器系统。
- 2 鉴于维护人道，以及捍卫国际人道法中保护平民和失去战斗力之战斗员的规则的道德考量，应禁止使用自主武器系统攻击人类。**要实现这一点，就最好禁止旨在或用于对人类使用武力的自主武器系统。
- 3 为了保护平民和民用物体，捍卫国际人道法规则并维护人道，对于未被禁止的自主武器系统，其设计和使用时应受到规制，**包括通过综合以下各项限制予以规范：
 - **对目标类型的限制**，例如仅限于攻打属于军事目标性质的物体；
 - **对使用时长、地理范围和规模的限制**，包括允许对特定攻击进行人为判断和控制；
 - **对使用情形的限制**，例如仅限于在不涉及平民或民用物体的情况下使用；
 - **对人机互动的要求**，特别是要确保有效的人类监督，以及及时的干预和停用。

上述建议与红十字国际委员会长期以来在为发展国际人道法，包括针对武器的禁令和限制标准做好准备这方面所发挥的作用相一致，¹⁸旨在维护人道原则，加强国际人道法，以应对将科技发展应用于自主武器系统这一作战手段和方法时所带来的挑战。

红十字国际委员会认为，现有的国际人道法规则并不能完全解决自主武器系统所引发的人道、法律和道德关切。需要新的规则来澄清、阐明国际人道法如何适用于自主武器系统，应对更广泛的人道风险和根本道德关切。而具有法律约束力的新规则具备法律确定性和稳定性的优势。红十字国际委员会担心，如果不设立此类规则，自主武器系统在设计和使用层面的进一步发展可能会损害国际人道法和人道原则目前为战争受害者所赋予的保护。

红十字国际委员会向所有国家提出建议，一方面旨在支持国家政策的制定，另一方面是支持国际上目前为应对自主武器系统带来的风险所开展的工作，包括《特定常规武器公约》政府专家组为商定自主武器系统规范和行动框架多方面问题而进行的工作。

令红十字国际委员会感到鼓舞的是，许多国家认识到需要对自主武器系统施以国际限制，而且已经在要求制定具有法律约束力的新规则，其他国家则更普遍地要求按照委员会提议的类似建议，确立国际公认的限制标准。委员会还承认，对于在何种程度上以何种形式对自主武器系统施加限制，目前仍存在分歧，而且一些国家认为国内措施就足以应对自主武器系统的问题。

在这一背景下，红十字国际委员会意在通过上述建议，推动树立共识，并促进各方取得进展，最终针对自主武器系统确立国际公认的有效限制标准。委员会期待与各国就上述建议进行进一步探讨，包括详细说明拟议的各项禁令和规定的范围中所涵盖的内容。

18 K. Lawand and I. Robinson, "Development of Treaties Limiting or Prohibiting the Use of Certain Weapons: The Role of the International Committee of the Red Cross" in R. Geiß, A. Zimmermann and S. Haumer (eds), *Humanizing the Laws of War: The Red Cross and the Development of International Humanitarian Law*, Cambridge University Press, June 2017, pp. 141–184.

红十字国际委员会将在其职责与专业范围内，继续与所有关注这一问题的利益相关方进行接触，并支持发表政治宣言、制定共同政策标准，或提供良好实践指南等旨在促进对自主武器系统施加限制，从而有效、及时应对此类武器所引发关切的工作，包括在《特定常规武器公约》的框架内就规范和行动框架的各个方面达成一致的工作。为此，红十字国际委员会随时准备与政府、武装部队、以及科技界和产业界代表等国际和国家层面的利益相关方开展合作。

ICRC POSITION ON AUTONOMOUS WEAPON SYSTEMS

使 命

红十字国际委员会(ICRC)是一个公正、中立和独立的组织，其特有的人道使命是保护武装冲突和其他暴力局势受难者的生命与尊严，并向他们提供援助。红十字国际委员会还通过推广和加强人道法与普遍人道原则，尽力防止苦难发生。红十字国际委员会创建于1863年，它是《日内瓦公约》和国际红十字与红新月运动的发起者。该组织负责指导和协调国际红十字与红新月运动在武装冲突和其他暴力局势中开展的国际行动。



红十字国际委员会东亚地区代表处
中国北京市建国门外大街9号齐家园外交公寓3-2
邮编：100600
电话：+86 10 8532 3290
传真：+86 10 6532 0633
邮箱：bej_beijing@icrc.org www.icrc.org
© ICRC, 07.2021



ICRC
微信



ICRC
微博