INTERNATIONAL REVIEW of the Red Cross

决定日益自主化军事系统之法律影响的因素

蒂姆·麦克法兰* 著
李强** 译

摘 要

本文探讨了决定使用自主军事系统之法律影响的五大因素。本文主张，带来法律挑战的系统是那些经编程可以“作出决策”的系统，而这些决策行为是受法律规制的。这种做法将决策控制权和决策责任从传统的决策者转移到负责开发和部署该系统的人员身上。本文还指出，由于士兵与自主性日益增强的武器系统之间的关系已发生改变，国际人道法规则对这一关系的调整作用可能在某种程度上具有局限性。

关键词：国际人道法；自主武器系统；武器发展；武器法；军事技术

---

* 蒂姆·麦克法兰是墨尔本大学法学院亚太军事法中心博士研究生。他正在跟随一个致力于“新兴作战技术对武装冲突法的挑战”项目的研究团队从事其博士研究。作者感谢其导师蒂姆·麦考马克教授和雷恩·利维耶博士，以及空军上校伊恩·亨德森先生，感谢他们对本文各稿给予的有见地的评论。文本的研究还得到还得到澳大利亚研究理事会探索项目资助计划的支持

** 李强，法学博士，中国政法大学副教授，军事法研究所副所长，主要研究方向为国际人道法和国际刑法。

引言

众所周知，各国都在积极发展军事系统，将利用先进的技术以帮助、补充以及某种程度上，代替人类士兵的战斗角色。如今进行中研发工作已制造出能复制战斗机飞行员[1]和哨兵[2]某些职能的机器，而且军事系统能力扩展至传统上人类操作者占主导地位的领域似乎是不可避免的。这些系统，一般被称为“自主系统”[3]，在接下来几十年内有可能给敌对行动在作战层面带来巨大变化。因此，这一前景引起了武装冲突领域律师的兴趣。

如果对这些拟议系统进行初步的法律分析，就会发现除了获取有关军事发展的可靠信息这种一般挑战外，那些试图调和这些系统与国际人道法的人们遇到了两项重要挑战。第一，这个问题的技术性很强，许多技术的性质都可能带来法律问题，尤其是机器人和人工智能方面的技术。除了远远超出律师的专业知识的专业问题外，这些领域内的尖端技术发展迅速，因此任何具体的法律分析可能尚未完成就已经过时。第二，少有例外，目前使用的系统看起来自主操作的能力很低，一般不会引起重要的法律问题。已公布的计划仅大致描述了更先进系统的未来形式、推动这些系统的技术及其使用方式。如今可用于检测的机器，例如密集阵近程防御武器系统[4]和采用主动终端制导的导弹以及类似的技术[5]，只体现了未来自主系统将具有的能力的初期形式。

---

1 Lockheed Martin, “UCLASS”, available at: www.lockheedmartin.com.au/us/products/uclass.html (all internet references were accessed in May 2015).

2 Jean Kumagai, “A Robotic Sentry for Korea’s Demilitarized Zone”, *IEEE Spectrum*, 1 March 2007, available at: http://spectrum.ieee.org/robotics/military-robots/a-robotic-sentry-for-koreas-demilitarizedzone.

3 本文不关注目前用于各种冲突中的无人驾驶飞机或无人机等遥控武器。此类设备的关键功能都是由操作人员手动控制的。虽然自主运载工具通常也是无人驾驶，但是本文只讨论由机器自主“作出决策”的能力引发的问题。

4 Raytheon, “Phalanx Close-In Weapon System (CIWS)”, available at: www.raytheon.com/capabilities/products/phalanx/.

5 For an overview of these technologies see Advisory Group for Aerospace Research and Development, North Atlantic Treaty Organization (NATO), *Precision Terminal Guidance for Munitions*, Advisory Report No. AGARD-AR-342, February 1997, available at: www.cso.nato.int/Pubs/rdp.asp?RDP=AGARD-AR-342.

由于高度自主的系统的最终能力并不清楚，对于它们是否符合现有国际人道法，有许多重大的问题需要回答。这些问题应尽快予以解答; 可以预见到的敌对行动的变化之大，至少需要进行预防性调查，而且它时间上的不确定性也带来了一定程度的紧迫性。此外，这对于律师来说是个不常有的机会，他们能够为新式军事技术的到来做准备，甚或在新系统部署前就影响其发展，而不是在滥用行为发生后再试图解决。

本文为系统分析由于使用自主军事系统而在国际人道法下可能出现的问题提供了基础。本文并未尝试进行完整的法律分析，而是探讨了机器自主性中最有可能具有重大影响的一些方面，提出了确定潜在法律问题的一些指导方针，以便为人们提供详细调查每个问题的基础。这种方法的基本特征是将法律分析的重点放在研发与操作自主系统的人员的决策和行动上，而不是机器的运作能力上。有两个术语方面的选择需要进行解释。

首先，虽然大多数讨论和最有争议的法律问题具体涉及自主武器，但本文通常更多地提到“自主军事系统”。就此而言，“系统”可以指任何执行指定任务的硬件或软件，或任何由硬件和/或软件组件构成的装置。无人操作的飞机或武器如炮塔属于本文所讨论的系统，此外还包括较广范围分布的传感器网络，“网络武器”，[6]甚或是为有助于指挥官决策而对数据进行分析的软件。“系统”可以大致被视为一种“作战手段”[7]，因为该用语出现在各种法律文书中。[8] 如下文所述，选择这一术语的原因是，相关技术可能被整合到军事硬件和软件中，这些硬软件通常不会被归类为武器，但仍然可能以与武器的自主能力相似的方式影响作战行动。例如自主情报收集、监控和侦察 (ISR) 系统，它使用安装在无人机、卫星、船舶或其他平台上的传感器收集和处理有关潜在目标的信息，之后再提供给人类武器操作员，这一系

6 网络武器是“用于，设计用于或旨在用于”实施“根据合理预期将导致人员伤亡或物体损毁的进攻性或防御性网络行动”的软件和/或硬件：Michael N. Schmitt (ed.), *Tallinn Manual on the International Law Applicable to Cyber Warfare*, Cambridge University Press, Cambridge, 2013, pp. 106 (Rule 30), 141 (Rule 41).

7 自主军事系统的使用也可能与作战方法有关，例如该系统使用的决策过程是最为重要的问题时，或者系统具有多种操作模式时。关于第二种情况，见：*ibid.*, p. 142, paras 4 and 5.

8 例如见《1949年8月12日日内瓦四公约关于保护国际性武装冲突受难者的附加议定书》，1125 UNTS 3，1977年6月8日 (1978年12月7日生效) (第一附加议定书)，第三部。

统在作出开火决定时所发挥的作用，类似于构成武器系统的一个自主决策系统的作用。

第二，与现在普遍的做法相反，本文不区分“自动”和“自主”系统。“自动”系统通常是指，在没有外部帮助的情况下执行低级任务，或者经编程按照固定程序严格执行重复性操作的系统；而“自主”系统是指，可以执行更高层次的指示，并且似乎在确定如何执行这些指令方面有一定程度的“选择”或“决策”能力。[9]本文认为，这种区分是人为的，定义不明确，在法律分析中毫无益处。各个系统的准确功能无法明确地归为明显不同的类别；相反，它们似乎存在于几个维度的连续体中，并且必须预料到，随着相关技术的进一步发展，系统能力也将超出现有定义的范围。律师如果希望提出具有持久性的法律原则，则应避免将该原则固定在可能发生变化的类别中，而这里提出的方法可以避免这一错误。此外，这种区分依赖于拟人化的概念，即机器作出决策，行使自由裁量权，并且某种意义上不是仅执行程序。虽然这些术语用于技术目的是有用的，但从法律角度来看，这些术语不是对有关系统的实际操作的准确描述，若用于法律讨论中，这些术语是具有误导性的。可以说，无论称之为“自动”系统还是“自主”系统，他们仍只是机器，其制造或编程的目的是为了减轻人类操作者手动执行一些决定和动作的负担。这一点将在下面更详细地讨论。

这里提出的关键结论是，律师评估“自主”系统时，不应根据它们能否在某种程度独立行事，或者能否表现出一些类似人的行为，而应根据授权其作出的决策性质以及人类操作员与系统的关系。在这方面，律师应该感兴趣的系统是那些可以代替人类作出某些受法律规制的决策或行动的系统，在这一过程中，将这些决策控制权和决策责任从手工操作员转移到另一方，也许是决定该系统行为或负责其使用的一方。

本文包含三个部分。首先讨论与拟议的军事系统的法律分析相关的机器自主性的三个方面。第二部分讨论了军事组织提出的发展建议的两个方面。

---

9 然而，关于其确切定义，即使在技术界中也存在着巨大分歧；see, e.g., the discussion in M. Shane Riza, *Killing Without Heart: Limits on Robotic Warfare in an Age of Persistent Conflict*, Potomac Books, Washington, DC, 2013, p. 13.

第三部分解释拟议发展的军事系统与国际人道法的关系，并为分析具体法律问题的人员提供指导方针：必须牢记作出法律上重要决策的责任仍然在人类身上，还必须了解，在使用自主系统时，这些决策的性质会发生变化，并且要谨慎区分法律和技术问题。

与法律分析有关的机器自主性问题

要对武器的“自主性”产生的影响进行合理评估，必须清楚了解机器的自主性的实际意义。就法律分析而言，没有必要深入研究实现自主行为的技术手段，更重要的是准确说明自主能力如何影响这类系统与整个世界之间的相互作用关系，尤其是它与操作员、主管以及受其影响的对象之间的关系。

不幸的是，自主性可能看似是一个明确的概念，但它却难以通过一个定义充分体现出来。[10]事实上，这个概念在不同的研究领域有着非常不同的含义。在这里，它在严格的技术意义上被使用，但即使在关于机器自主性的技术文献中，似乎每一个作者都有自己的定义。然而，在高层次上，我们可以说，大多数定义主要集中在自主性的两个相互关联的方面。一些定义侧重于操作员与系统之间的交互：自主系统是可以“在长时间内无任何形式的外部控制的情况下”进行运行的。[11]其他的定义则从系统自身的能力来表述：自主是“系统在限制范围内选择和决定自己的行为的能力”。[12]当然，这些和其他定义仅涵盖机器自主性有关特定目的一些方面，希望进行调查的律师还必须决定哪些方面与调查有关。本节并不旨在提供另一个 (难免不全面的) 定

10 See, e.g., the discussion in Henry Hexmoor, Christiano Castelfranchi and Rino Falcone, “A Prospectus on Agent Autonomy”, in Henry Hexmoor, Christiano Castelfranchi and Rino Falcone (eds), *Agent Autonomy*, Kluwer, Boston, MA, 2003, p. 3.

11 George A. Bekey, *Autonomous Robots: From Biological Inspiration to Implementation and Control*, MIT Press, Cambridge, MA, 2005, p. 1. Similarly, “[a] system with a high level of autonomy is one that can be neglected for a long period of time without interaction”: Michael A. Goodrich and Alan C. Schultz, “Human-Robot Interaction: A Survey”, *Foundations and Trends in Human-Computer Interaction*, Vol. 1, No. 3, 2007, p. 217.

12 Charles François (ed.), *International Encyclopedia of Systems and Cybernetics*, Vol. 1, K. G. Saur, Munich, 2004, p. 51. Similarly, “[a]utonomy is a capability (or a set of capabilities) that enables a particular action of a system to be automatic or, within programmed boundaries, ‘self-governing’”: Defense Science Board, *The Role of Autonomy in DoD Systems*, US Department of Defense, July 2012, p. 1, available at: www.acq.osd.mil/dsb/reports/AutonomyReport.pdf.

义，而是力图通过提出在实践中如何实现机器自主性的高层次观点来指导读者，并讨论了自主系统在军事上可能具有最大法律意义的一些方面，律师研究发展建议应该考虑到这一点。特别是，如下所述，不应仅将这一过程视为武器发展过程。[13]

如何实现自主?

为了分析，首先必须对自主武器系统的运作原理有基本了解。设计师通常将自主系统介绍为由两个主要部分构成："工厂"是将予控制的系统或流程，"控制器"是直接支配工厂行为的设备。[14]"工厂"一词来自化学工程；在军事方面，它将指的是，不具备自主操作能力的设备由人员直接操作，例如车辆或炮塔。自主系统的控制器由按照开发者编写的程序管理车辆，武器或其他设备的硬件和软件组成。[15]图1和图2从概念上分别概述了这些组件在手动操作的系统和自主系统中如何协同工作。实线箭头显示了各种组件之间的典型相互作用。这些图并不涉及具体的武器系统；它们仅仅描述了每种类型的系统的一般功能。

为了进行法律分析，现在可以对机器的自主性发表以下几点看法。

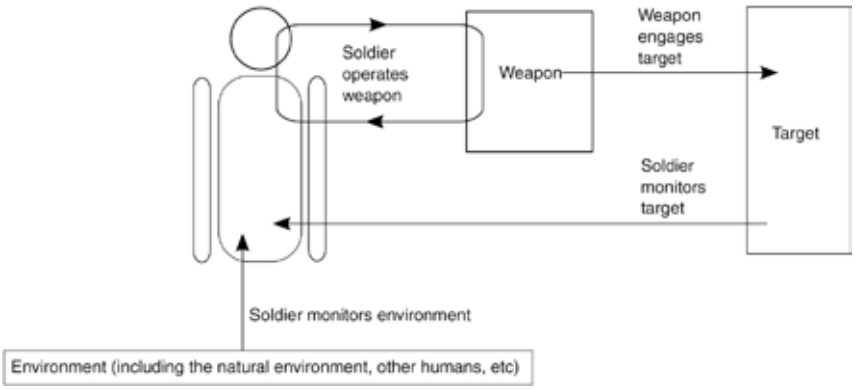

Figure 1. Manual weapon system.

13 关于将自主和遥控武器系统归类为武器是否适当的讨论，见廖显恩撰写的《自主和遥控武器系统的分类与合法性》，*International Review of the Red Cross*, Vol. 94, No. 886, 2012, p. 627.

14 See, e.g., Zdzislaw Bubnicki, *Modern Control Theory*, Springer, Berlin and New York, 2005, pp. 3–4.

15 关于自主控制系统的一般性介绍，see Panos J. Antsaklis, Kevin M. Passino and S. J. Wang,"An Introduction to Autonomous Control Systems", *IEEE Control Systems*, Vol. 11, No. 4, 1991, p. 5.

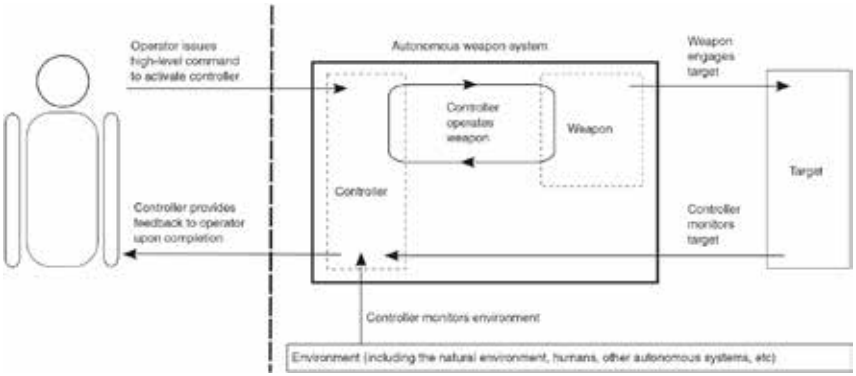

Figure 2. Autonomous weapon system.

自主性既不是固定的、也不是统一的

法律学者通常将自主武器称为一种独立类别的设备，这些设备与非自主系统很容易区分，并且还会提到特定系统所表现出的自主“级别”，仿佛这些级别是系统的固有属性。[16]这种简单的区别并未反映现有系统的能力范围，并且与各种武装力量已经公布的发展路线图[17]不符。事实上，拟议军事系统所展现的自主水平可能会以复杂的方式发生变化。

可能存在的自主性程度也差别很大，操作员和自主系统之间的任务分配方式也可能有所不同，系统的行为可能会根据正在执行的具体任务和系统运作环境而发生改变。就某一行动而言，出于法律或其他目的，要确定人类操作员和电脑控制系统所实施的相对控制程度可能都是一个复杂的过程。本节的简要说明并不旨在为读者提供所需的所有技术知识，而是展现在人类与先进自主系统交互的环境中对任务实施的控制具有流动性。已经提出了描述这个连续体的许多方法。[18]为人熟知的一个方法提出了自动化的十个级别，见

16 See, e.g., Markus Wagner, “Taking Humans Out of the Loop: Implications for International Humanitarian Law”, *Journal of Law, Information and Science*, Vol. 21, No. 2, 2011.

17 See, e.g., Army Capabilities Integration Center, US Army, *Robotics Strategy White Paper*, 19 March 2009, available at: www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA496734; US Air Force, *Unmanned Aircraft Systems Flight Plan 2009–2047*, 18 May 2009, available at: www.fas.org/irp/program/collect/uas_2009.pdf.

18 Peter A. Hancock and Stephen F. Scallen, “Allocating Functions in Human–Machine Systems”, in Robert R. Hoffman, Michael F. Sherrick and Joel S. Warm (eds), *Viewing Psychology as a Whole: The Integrative Science of William N Dember*, American Psychological Association, Washington, DC, 1998, p. 521.

表1。不过，这些“等级”只是描述连续体上节点的方法；在现实中，机器的自主性并没有明确的级别。

计算机控制的程度

机器的自主性也是有程度可言，从对某些功能的完全人为控制状态到完全机器控制的范围内，自主性逐渐变化。

美国国家标准与技术研究所开发了一种更为复杂的分类法，即无人系统自主水平框架。[19] 该框架将自主性水平的特征归纳为一个三维模型，即任务复杂性、环境复杂性和免受人类干预的独立性这三个维度，其中的每个因素又进一步分为一系列较低级别的考量因素。

表1：Sheridan和Verplank的自动化十个级别

100%人类控制	1. 人类完成整个工作，直至将其转变为电脑来实现。 2. 计算机帮助确定方案。 3. 计算机帮助确定方案并提出方案，无需人类遵循。 4. 计算机选定行动，人类可以、也可以不做。 5. 计算机选定行动，如果人类批准则执行。 6. 计算机选定行动，告知人类并留有充足时间停止该行动。 7. 计算机完成全部工作，并将其所为告知人类。 8. 计算机完成全部工作，但只有在人类明确要求时才将其 所为告知人类。 9. 计算机完成全部工作，但只有在电脑决定它应告知人类时，才告知人类[20]。
100%计算机控制	10. 如果计算机根据自身的决定，完成了全部工作，而且只有当它决定应当告知人类时，才会告知人类。

Source: Thomas B. Sheridan and William L. Verplank, Human and Computer Control of Undersea

19 NIST Engineering Laboratory, National Institute of Standards and Technology, *Autonomy Levels for Unmanned Systems*, 16 June 2010, available at: www.nist.gov/el/isd/ks/autonomy_levels.cfm.

20 此处的“决定”指的并不是类似于人的决策能力；它仅指计算机是按照其程序而不是人类操作员的指令来开始实施行动的。

此外，其他组织还提出了用于不同目的的自主级别的分类法，例如美国航空航天局有关航天器任务评估和重新规划工具[21]的八级模型，以及美国陆军科学委员会从其对人机界面问题中研究出来的十一级模式。[22]

没有必要深入研究这些相当详细的模型。重要的一点是，试图将某种机器描述为自主、半自主或手动操作是没有意义的，因为这些并不是相互独立、客观存在的类别。自主能力是不断变化的现象。

人机交互和任务分配方法

即使是在完成分配任务时无需要大量的人力投入的高度自主系统，也会与其他实体（人力和电子）一起运行，并且需要与这些实体交换信息，以便接收指示、协调工作、报告结果等等。人机交互研究以及最近人与机器人的交互研究是多学科的领域，在过去二十年中随着计算和机器人功能的扩大，受到越来越多的关注。[23]出现了其他几种范例来描述这种相互作用可能以怎样的结构出现，怎样将任务分配给任何一个实体，并说明了对人类作用的影响。从广义上说，在使用自主系统时，人类可能会担当监督或协作角色，并且可能在执行一系列任务期间在这些角色之间移动。例如，一个分类法列出了人工在与机器人工作时可能发挥的五个角色：主管、操作员、技工、同伴和旁观者。[24]这些角色可能不是静态的；有一个称为混合主动交互的设计范例“是指一种灵活的交互策略，每一方可以按其所长为任务做出贡献”。[25]在混合主动范式中，人类和计算机（或机器人）各自的角色往往并不是提前确定的，而是由不断变化的情况来决定。在不同阶段，人类操作员或机器在另一方协助的情况下可以对任务有直接的控制，或者它们可以独立工作。

---

21 Ryan W. Proud, Jeremy J. Hart and Richard B. Mrozinski, *Methods for Determining the Level of Autonomy to Design into a Human Spaceflight Vehicle: A Function Specific Approach*, NASA Johnson Space Center, September 2003, p. 4, available at: www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA515467.

22 US Army Science Board, *Ad-hoc Study on Human Robot Interface Issues*, September 2002, p. 16, available at: www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA411834.

23 For an overview, see M. Goodrich and A. Schultz, above note 11.

24 Jean Scholtz, “Theory and Evaluation of Human Robot Interactions”, in *Proceedings of the 36th Annual Hawaii International Conference on System Sciences*, Hawaii, 6–9 January 2003.

25 Marti A. Hearst, “Trends & Controversies: Mixed-Initiative Interaction”, *IEEE Intelligent Systems*, Vol. 14, No. 5, 1999, p. 14.

人机交互是一个活跃的研究领域，新原则不断涌现。虽然可以安全地假设人类整体上将继续对武装冲突中使用的机器人武器进行高度控制，但在评估特定个人的作用及其实施的控制程度时，律师仍应谨慎。

功能不同而引发的变化

人们通常将自主性视为某个系统的整体特点，但在实践中用于自主运行能力的技术将被、并且正在应用于构成军事硬件和软件部分的各个子系统。在机器人系统方面，发展分别在自主导航、自主攻击和先进武器系统所需的其他功能方面发生。未来几年会发展出自主运作能力更强的系统，没有理由认为其所有职能都将受到同等程度的人力监督；对于某些功能而言，自动化成本/效益分析较为有利的情况下，会比其他功能获得“更多自主性”。因此，系统同时进行不同任务时就可能具有多种“自主”水平，而研究此类系统的律师则需针对航线规划、导航或武器释放等有关功能，评估其行为以及人机参与程度。

例如，仔细研究澳大利亚国防军采用的六步攻击过程，[26]就会发现也许不能授权可自主定位，观察潜在目标并可采取一些预防措施以尽量减少附带损害的武器系统自主发射武器，可能需要人类操作者或其他自主系统来评估目标是否为有效军事目标，以及任何预期的附带损害是否与攻击的预期军事利益不成比例。[27]如果某一事故涉及此类系统，需要调查是否可能对平民实施了攻击，重要的是确定可能在哪个阶段出现了错误，以及在这一阶段执行了什么级别的人类和机器控制，以确定人类操作者是否实施了足够的控制，可被追究刑事责任。同样，对于“一个系统由多个系统构成”的情况，也需要有类似的考虑，该情况涉及具有不同自主程度的系统间的通讯。

26 Ian Henderson, *The Contemporary Law of Targeting*, Martinus Nijhoff, Boston, MA, and Leiden, 2009, p. 237.

27 关于攻击过程的不同阶段如何实现人类控制和自主运作之间的平衡的详细讨论，see Mark Roorda, “NATO’s Targeting Process: Ensuring Human Control Over and Lawful Use of “Autonomous” Weapons”, Amsterdam Center for International Law Research Paper No. 2015-06, 2015, available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2593697.

情形不同而引发的变化

自主系统的功能是指系统正在执行的任务，而情形是指执行这些任务时所处的状况。如上所述，即使设计者确定特定系统功能将受到计算机某种程度的控制，但控制程度可能根据情形而变化，而有关情形可以需要依据下列因素确定，例如：

- 任务阶段，如规划、启动、实施或终止；
- 干扰事件，例如与操作员失去联系的情况下，通常远程操作的无人机可能被编程为自主地返回基地、盘旋、甚至可实施自卫；
- 在任务期间出现了意想不到的机会。

美国国防部对这种变化性总结如下："关键是人类以及代替人类的计算机将在任务阶段和阶层进行主动权和角色交替，以便随着情况的变化而应对新出现的事件、干扰和机会。"[28]

总结

与人们对终结者这样的机器人通常具有的印象不同，机器的自主性是一个更加细微的现象。它是一系列迅速发展的技术，一些还处于起步阶段，无疑许多关键问题目前尚未出现，但有两个特别突出的问题尤其引起正在研究这一领域的律师的注意。首先，机器的自主能力主要直接影响其主管和操作者，而不一定 (在武器的情况下) 影响攻击对象。特别是，在许多重要方面，操作者的角色发生了改变，但并没有被消除; "完全"的自主性并不存在，没有完全无需人类干预而独立行事机器 。第二，操作员和机器之间的关系的确切形式，会因系统不同、阶段不同或任务不同而发生变化。律师一定要警惕那种简单将武器称为"自主"武器的观点。通过研究美国国防科学委员会最近提出的"自动系统参考框架"[29]，可以了解自主系统、自主系统操作者及其主管之间关系的可能复杂性，该框架在系统设计阶段将功能和职责分配给计算机或其中一个操作人员或主管。该框架考虑了自主性如何支持军事层级中特定系统的各种人类操作员，如何方便他们之间的沟通，分配给人或机器的任务在执行过程中可能发生怎样的变化，以及其他几种因素。

28 Defense Science Board, above note 12, p. 27.

29 *Ibid*., p. 24.

自主性是关于机器和操作员之间的关系

一般就机器人技术而言，特别是发展自主军事系统的提案，自主性是指能力，而不是特定技术，特定设备或某种行为。自主性是系统执行其功能的能力，与人类操作员的交互少于手动系统所需。因此，自主性关系到系统与人类操作者之间的关系，而不是系统任务的性质或者执行该任务的方式。如上所述，这种关系根据系统自主能力程度在一定范围内变化，从完全人为控制到机器对特定任务在实际上的完全控制。

当手动系统或过程被替换为能够进行一定程度的自主操作的系统时，控制器在一定程度上“成为”手动系统的操作者。操作员 (或系统开发人员) 对如何控制系统的理解体现在软件中，并编入控制器。操作者操纵系统的物理手段被转换成控制器可以激活的一组致动器。[30]还有某种形式的传感器或反馈机制，控制器可以通过该机制监视系统。此外，还有允许控制器监测相关环境因素的其它传感器。控制器根据其程序操纵来自所有传感器的信息，并产生发送到致动器以调节系统的输出信号。众所周知的“感知-思考-行动”范式概括了这一过程，这种范式通常被用作机器人的操作定义。[31]

军事发展提案通常从“观察，定向，决策和行动” (OODA) 循环过程的角度讨论自主性，[32]这是由美国空军上校约翰·博伊德提出的关于战斗员的决策循环周期模式。OODA模式描述了，为实现某些目标而对所处环境进行观察并应对变化时持续的心理和物理过程。高级目标可能涉及到几个子任务，每个子任务都有自己的OODA循环来完成，以追求总体目标。在手动系统中，循环中的所有步骤都由人完成：观察环境以提取原始信息，通过处理该信息来形成有用的模型来对自身进行定向，从而基于该模型作出决定并执行

30 致动器仅仅是控制器控制工厂的一个装置。一个例子是一种电动马达，它根据炮塔以软件为基础的控制系统发射的信号来转动炮塔。

31 “感知-思考-行动”范式指的是机器人感知其环境，使用感知到的信息根据其程序“作出决策”，并将这些决策付诸行动的一个持续进行的过程；see, e.g., G. A. Bekey, above note 11, p. 2.

32 例如见，美国空军，上文脚注17，第16页。OODA循环过程可以被视为“感知-思考-行动”过程的另一个表现形式，一些有关人机交互的文献中也提到了一种类似于OODA的循环过程，用以描述自主武器系统可以实施的四类功能：信息获取、信息分析、决策选择以及行动执行。See, e.g., Raja Parasuraman, Thomas B. Sheridan and Christopher D. Wickens, “A Model for Types and Levels of Human Interaction with Automation”, *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 30, No. 3, 2000, p. 288.

该决定。在OODA循环方面，开发自主系统的目的是将循环周期部分或全部分配给机器，以实现一些操作优势，例如提高速度或耐力，降低成本或对操作人员的生命危胁。高度自主的系统可以执行实现某些目标所需的OODA循环周期中大部分或全部过程，仅利用操作员的高级指令作为循环决策阶段的指导，该过程中“最接近”的人类职能类似于指挥官。自主性较低的系统只能执行低级循环，或仅执行循环的某些部分，并且必须与人类操作员一起实现高级目标，人类操作员更多发挥协作作用。这种系统的一个例子可能是雷达警报接收器; 更简单的系统也许仅可探测到可能的威胁并向操作人员发出警告，而更高度自主的系统可能无需人为干预即可实施对策，例如发射干扰物和曳光弹。

自主性的作用体现在操作员和机器之间的关系上，这一事实对律师有三个重要意义。

首先，自主系统将代替人类执行部分或全部任务，但执行方式不一定与人类不同。从机器自主概念本身推断不出，与人类或人类团队手动执行任务相比，自主系统必然以不同的方式执行相同任务。当然这并不是说，未来自主系统的功能将与人类操作的系统完全相同；无需持续人机互动的系统所具有的“持续性”，快速整合来自许多来源的数据的能力，与人工操作系统相比承受更大风险的能力以及其他特性将大大提高其性能。然而，这些差异虽然在运作上非常重要，但某种程度上并未触及自主问题的核心。例如，无人机已经拥有较高程度的持续性，却不一定具有与高度自主性相关的任何能力，而且不会引起相同的法律问题。远程操作系统和自主系统同时具有的这些能力常常导致人们对这两类机器之间的区别产生困惑，因为这两种机器能够以不同的方式“代替”人类，并且都会让人联想到没有人类参与的机器作战的情形。不过，律师必须清楚这两者之间的区别，因为他们会引发不同的法律问题上。今天的无人机和类似的装置可以扩展人类战斗员的作战范围和能力，以及自主系统也是如此，但自主系统还将改变导致武器激活的决策过程。

因此，对拟议自主系统进行法律分析必须基于这样一个前提，即系统的新颖性可能体现在系统上的武器被激活的过程中。仅在某系统具有某种自

主运作能力的基础上，就试图将特定的性能标准归因于该系统，这样做是徒劳无益的; 在此基础上可以确定的一点是，它将减少或消除人类操作员对系统部分或全部OODA循环的直接参与。将任务从人类重新分配到计算机的事实，并不一定会改变该任务的性能。

第二，如果将具体的决策和行动从人类操作者转移到自主系统，同时那些决策和行动是法律义务的对象，那么重要的是要考虑这些决策和行动如何归咎于负责人，以及这些负责人是否是现行法律规定的负责人。如果承担法律义务的操作员由自主系统协助或与自主系统协作，则必须思考该操作者是否仍然承担法律规定的作用。

第三，将自主系统描述为“无人控制”的“独立”[33]机器是不准确的；[34]人机关系仍然存在，仅是有所调整。 硬件和软件开发人员在设计阶段做出的选择将从此决定系统的行为。 任务规划者和其他人也将对每个任务施加限制；例如，在自主无人机的情况下，必须事先提交飞行计划，指定要覆盖的区域和任务的持续时间 (并且无人机必须遵守该计划)，同时关于携带多少燃料和哪些武器的决定，将进一步指导和限制行动中可能发生的情况。[35]在这些方面，尽管可能没有直接监督，但人类总是会提供一定程度的指导。

自主系统仍是机器

律师必须避免陷入自主系统拟人化的陷阱。如今的尖端武器系统只是执行编入软件的指令的机器，这是不存在争议的，本文认为设计者设想的未来高度自主的系统也仅仅如此。当然，它们在许多方面会更复杂，更有能力；它们可能能够利用来自当今系统范围以外的来源的信息，并在面对更多不确定性的情况下处理该信息，并且将在更混乱和艰巨的环境中有效运行。然而，这些增强的性能是由于系统所控制的硬软件方面的改进；它们不会从根

33 M. Wagner, above note 16, p. 159.

34 Gary E. Marchant et al.,“International Governance of Autonomous Military Robots”, *Columbia Science and Technology Law Review*, Vol. 12, 2011, p. 273.

35 William Boothby, “How Far Will the Law Allow Unmanned Targeting to Go?” in Dan Saxon (ed.), *International Humanitarian Law and the Changing Technology of War*, Martinus Nijhoff, Boston, MA, and Leiden, 2013, p. 56.

本上改变系统的性质，仍属于由计算机控制的机器。系统的能力和局限性仍将直接或间接地来自于人类决策和行动。

当今自动化武器系统中的基于软件的控制器基本上是用于特定目的的计算机，其运行的程序代替人类操作员控制武器。虽然这些控制器在设计和目的方面可能具有高度的专业性，但仍然是存储程序计算机的形式，是一类还包括诸如个人计算机之类的常见项目的设备。存储程序计算机的决定性特征是由程序员输入的指令存储在机器的存储器中，并用于管理其操作。[36]除非发生重大的技术转变，未来的自主系统仍将采用基本相同的技术；系统仍将由人类开发人员编写的软件控制。

在关于尖端武器系统的讨论中，往往忽略了即使非常复杂的程序也只是一组预定义的指令这一事实。观察者并不总是很清楚复杂的机器只是执行指令而不是独立运行。即使是仅具有低级功能的系统也通常由指令驱动，通过“如果 <X发生> 那么<执行行动A> ，若未发生则 <执行行动B>”的形式。看似是系统本身在两个不同的行动方案之间进行“选择”，但实际上是由编程人员预先作出的选择；该程序员的意志被编入系统的存储器中，在检测到先前确定的触发机制时得以体现出来。例如，如果假设一个自主无人机具有相机和图像识别程序，可以将其视野内的人与已知叛乱分子图像的数据库相匹配，那么类似于“如果<相机图像与数据库图像匹配度超过95%>则<瞄准并攻击>若非<继续搜索>”的一个指令，看似是无人机本身选择了目标，但实际上是系统开发人员事先选定了目标以及攻击条件。

这里，之所以提到计算技术，是因为律师们应当清楚，并没有具有“选择”或能“真正自主”开展行动的“智能机器”，这非常重要。[37]没有哪一台计算机能够自己选择运行存储在其存储器中的程序，或者对程序中是否执行特定指令行使裁量权；任何看似是计算机作出的“选择”只不过是编入软件的其他指令的结果。从根本上说，计算机的唯一功能是运行其上安装的任何软件。

36 William Aspray, “The Stored Program Concept”, *IEEE Spectrum*, Vol. 27, No. 9, 1990, p. 51.

37 See, e.g., Chantal Grut, “The Challenge of Autonomous Lethal Robotics to International Humanitarian Law”, *Journal of Conflict and Security Law*, Vol. 18, No. 1, 2013, p. 5.

如果人们能看到计算机执行的是由开发人员编入软件的低级明示指令，则很容易理解计算机不在任何重要的法律意义上独立行事，但是当具有更先进能力的系统被“训练”去实施一些与发射武器等重大行动有关的行为时，这一点就未必如此显而易见了。[38]类似的挑战还会出现在操作人员或指挥官仅给自主系统制定一个高级别任务目标，自主系统自行制定一系列低级子任务的情况下。当然，人工智能和机器学习是复杂的领域，这些领域远远超出了本文的范围[39]，但是为了进行法律分析，不必深入研究具体算法的细节。但是，律师们应该认识到，人工智能系统的发展实际上只是软件开发中的一个活动，与开发一个更简单的低级程序时进行的人类活动没有区别。在这两种情况下，开发人员都会为系统设计一些所需的行为，并编写程序，将该行为编入系统中。区别在于程序中使用的算法：简单程序开发者以 “如果<目标符合这些参数> 则 <开火>”的形式直接为行为编码，而 “智能”机器开发人员所编写程序的功能，则是为应对执行任务期间出现的环境刺激，形成一系列最佳行动。在某种意义上，开发者和武器发射动作之间又有一层抽象关系，就是开发人员可能不了解促成武器发射的一系列具体事件，但这一系列事件很有可能源于一些数据，而该系统就这些数据进行过训练;尽管如此，运行程序的最终结果仍然是由一项规则告诉系统发射武器，就像一个由固定规则组成的更简单的程序一样。这另一层抽象关系导致将特定结果与特定人类指令相匹配的过程变得复杂，但并不会改变计算机只执行其开发人员制定的指令这一事实。

要明确的是，“智能”系统自行做出“决定”的能力这一说法具有误导性。虽然可能并未对这类系统会遇到的每一种情形，预先编写准确的应对措施，但有些系统经编程可发展出应对措施，因此他们仍根据其编程来操作，无论其实际采用的具体行为能否在开发期间或部署时被预见到。这样的机器仍然只是人类 (开发人员以及某些情况下为其使用负责的人) 意志的工具；准

38 在此语境下，训练指的是给人工智能系统提供多组示例数据，这些数据代表其将面临的任务和正确的应对方式，以便在这些任务中诱导产生最佳结果的行为。训练本质上是归纳性的，因为受过训练的系统遇到的“真实”情况在某些方面将不可避免地与培训的例子不同，因此容易产生错误。

39 For a general overview see, e.g., Stuart Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach, Prentice Hall, Englewood Cliffs, NJ, 2009*.

确地说，它并未一个独立的决策者。例如，具有自主航行能力的无人机或许能够在飞行中应对其软件未具体规定的事件，但是他们仍然根据编程的规则来应对意外事件。行为并不源于机器本身，而是源于其开发人员的思想。

军事发展提案中有关法律分析的几个方面

自主性将扩展至武器系统之外

正如引言中所述，本文讨论自主军事系统，而不仅仅是武器。即使在专门针对国际人道法的分析中，有必要说明系统自主能力的影响，这些系统本身可能不会发生任何敌对行为，但仍可能对具有法律后果的决定或行动产生一些影响。这种能力的主要例证就是自主ISR系统，该系统对潜在目标实施定位，识别和跟踪。例如，美国国防部最近宣布了自主研究试点计划(ARPI)，该计划“旨在促进发展创新性、跨学科科技的发展，推动自主系统满足未来DOD系统和任务的要求”。[40] ARPI 邀请提案将ISR视为其技术挑战领域之一：

通过增加ISR平台本身的机器感知，推理和智能水平，可以实现更有效的工作负载平衡。这包括ISR资产的管理和闭环控制，以适应其环境和任务情况，以收集适当的相关数据。[41]

这方面的初步行动已经发生了；在美国，国防高级研究计划署从事军事成像和监视技术 (MIST) 计划，其目标是“开发一种全新的光学ISR能力，可以生成高分辨率的三维图像，在远远超过现有光学系统可能的范围内定位和识别目标”。[42]根据该计划开发的系统将能够自动辨认和识别潜在目标。[43]

40 Department of Defense, *Autonomy Research Pilot Initiative (ARPI) Invitation for Proposals*, November 2012, p. 1, available at: www.auvac.org/uploads/publication_pdf/Autonomy%20Research%20Pilot%20Initiative.pdf.

41 *Ibid*., p. 4.

42 See: www.darpa.mil/program/military-imaging-and-surveillance-technology.

43 Strategic Technology Office, DARPA, “Military Imaging and Surveillance Technology – Long Range(MIST-LR) Phase 2”, Broad Agency Announcement No. DARPA-BAA-13-27, 12 March 2013, p. 6, available at: www.fbo.gov/index?s=opportunity&mode=form&id=78b0ddbf382678fa9ace985380108f89&tab=core&_cview=0.

作为ARPI的一部分，追求这一研究路线的动机是，如今战场上的情报、监视和侦察数据大大增加，这是前所未见的。处理，开发和传播数据的分析人员可能无法在试图跟踪目标，推断出来源并提供分析反馈 (实时或事后分析) 的同时，对各种数据输入 (图像，视频，通信和人类ISR数据) 进行整合和分析。[44]

可以推断，“智能”提升了ISR平台将用于分析某些原始数据，然后再将结果传达给人类，从而使得ISR系统能够影响操作人员对战场的认识。这种分析传感器数据的方式会明显作用于先进ISR系统对决策的影响，有关发射武器或执行其他具有重大法律后果的行动的决策。

有些不参与战斗行动、但仍然对结果具有某些外围影响的武器是自主车辆，它们能够运载货物、用品甚至人员。为实现自动补给行动自动化正开展着各项工作，例如美国海军研究自动航空货运/公用事业系统 (AACUS) 计划。[45]随着在阿富汗的K-MAX无人值班货物补给直升机[46]的成功，AACUS计划的目标是“开发先进的自主能力，利用无人和可有人驾驶的垂直起降 (VTOL) 系统，以实现快速货物交付”。[47]该计划将生产一个系统，可安装在适合的飞机上，以响应部署部队的需求，自动 (但也受某种监督) 规划路线，避开障碍物和恶劣天气，选择合适的着陆点提供物资，并撤离伤亡人员。[48]尽管由类似AACUS的系统所控制的车辆不会携带武器，但它们仍然是大型、快速移动的物体，必须在人和其他车辆附近运作并且可能携带有害物质或带来其他危险。[49]因此，如果用它们来运送伤员或其他一些情形，就会产生一些法律风险。

---

44 Department of Defense, above note 40, p. 4.

45 Office of Naval Research, “Autonomous Aerial Cargo/Utility System Program”, 27 September 2013, available at: www.onr.navy.mil/en/Science-Technology/Departments/Code-35/All-Programs/aerospaceresearch-351/Autonomous-Aerial-Cargo-Utility-AACUS.aspx.

46 Lockheed Martin Corporation, “K-MAX”, available at: www.lockheedmartin.com/us/products/kmax.html.

47 Mary Cummings and Angelo Collins, Autonomous Aerial Cargo/Utility System (AACUS): *Concept of Operations*, Office of Naval Research, p. 2, available at: www.onr.navy.mil/~/media/Files/Funding-Announcements/BAA/2012/12-004-CONOPS.ashx.

48 *Ibid*.

49 Anthony Finn and Steve Scheding, *Developments and Challenges for Autonomous Unmanned Vehicles: A Compendium*, Springer, Berlin, 2010, p. 156.

因此，作战行动所使用系统中出现的自主能力，要求在广义上理解自主性，而不仅只是一种新型武器的一个性质。另外，如下一节所述，AACUS和类似系统直接与其他无人系统[50]进行谈判的能力，可能会使识别哪些系统与特定调查相关的任务变得复杂。

自主系统间的协作

如上所述，自主系统改变了武器和操作者之间的关系。同样重要的是名义上独立的自主系统之间的关系。军事和政府组织发布的大多数发展路线图明确表明，自主系统必须相互配合。[51]这种互操作性可能采取一系列形式。

许多目前的军事和民用研究项目都集中在分散型合作机器人群体的行为上，俗称“蜂群”，它们作为一个单一的系统来追求一些目标。其中一个例子是美国陆军微型自主系统和技术 (MAST) 计划[52]，旨在创建“配备微型传感器的各种自主移动平台构成的系统，可快速、安静、可靠地探索和寻找目标”；[53]即小型空中和地面车队，士兵将其用于探索和绘制诸如城市环境或洞穴等复杂环境。为集群装备武装的可能性也有所提高。[54]

美国国防高级研究计划署的海德拉发展项目则使用了自主系统间的另一种协作形式。[55] 海德拉方案“旨在开发一个分布广泛的由无人载荷和平台构成的海底网络”[56]，本质上是一个无人潜艇平台的系统，可用于发射各种无人机和无人海底车，靠近敌方行动。尽管还处于初期阶段，该计划设想了潜艇“母舰”与其无人机和无人海底车有效载荷直接相互作用的系统; 例如，

50 M. Cummings and A. Collins, above note 47, p. 2.

51 See, e.g., US Department of Defense, *Unmanned Systems Integrated Roadmap FY2011-2036*, No. 11-S-3613, 2011, pp. 49–50.

52 US Army Research Laboratory, “Micro Autonomous Systems and Technology (MAST)”, 25 February 2011, available at: www.arl.army.mil/www/default.cfm?page=332.

53 MAST, “Research Thrusts”, available at: www.mast-cta.org.

54 Peter W. Singer, *Wired for War: The Robotics Revolution and Conflict in the Twenty-First Century*, Penguin Press, New York, 2009, p. 229.

55 Federal Business Opportunities, “Hydra”, 22 August 2013, available at: www.fbo.gov/index? s=opportunity &mode=form&id=4cc32f06144bd6f3eba18655135d6155&tab=core&_cview=1.

56 DARPA Tactical Technology Office, “Hydra”, available at: www.darpa.mil/Our_Work/TTO/Programs/Hydra.aspx.

无人海底车将与母舰对接并进行装备，收集情报信息以供其任务使用，之后将在任务中获取的信息传递母舰，再送到当局。[57]

这些和类似项目，以及与之相关的战术和战略，是一个总体大趋势的一部分，将军事力量中相互分离的部分发展为紧密地相互连接在一起。也许这种趋势最广为人知的表现形式，是以网络为中心的战争 (NCW) 的教义[58]，这是实施敌对行动的一种方式，它强调武装力量元素之间共享信息的重要性，以便最好地利用这些信息和武装力量的能力。它最初是在美国发展起来的，但北约[59]以及澳大利亚，[60]英国[61]和其他国家正在追求或至少提出类似的军事系统互联的方法。这种信息共享的预期好处之一是一定程度的分散操作或“自我同步”的可能性，其中两个或多个实体可以直接交互并协调其努力，而不采用传统的分级命令和控制结构。[62]相关地，其中一些综合系统具有显著的自主操作水平，信息的共享不一定由人进行或直接监督，因此，例如，如果一个自主系统收集的情报被另一个自主系统用于进攻行动，那么所有涉及到的系统都可能成为法律调查的一部分。

虽然NCW等概念本质上与军事系统自主能力的发展无关，但它们提供了这种先进系统可能运行的环境类型以及可能成为常态的军事系统之间的一体化

57 John Keller, “DARPA Considers Unmanned Submersible Mothership Designed to Deploy UAVs and UUVs”, *Military & Aerospace Electronics*, 23 July 2013, available at: www.militaryaerospace.com/articles/2013/07/darpa-uuv-mothership.html.

58 See, generally, David S. Alberts, John J. Garstka and Frederick P. Stein, *Network Centric Warfare*, 2nd revised ed., Department of Defense Command and Control Research Program, 1999.

59 NATO, “NATO Network Enabled Capability”, 27 October 2010, available at: www.nato.int/cps/de/SID-815535E4-57782C82/natolive/topics_54644.htm.

60 M. P. Fewell and Mark G. Hazen, *Network-Centric Warfare: Its Nature and Modelling*, Defence Science and Technology Organisation, September 2003, available at: http://dspace.dsto.defence.gov.au/dspace/bitstream/1947/3310/1/DSTO-RR-0262%20PR.pdf.

61 UK Ministry of Defence, *Network Enabled Capability*, JSP 777, 2005, available at: http://webarchive.nationalarchives.gov.uk/20121026065214/http://www.mod.uk/NR/rdonlyres/E1403E7F-96FA-4550- AE14-4C7FF610FE3E/0/nec_jsp777.pdf.

62 D. S. Alberts, J. J. Garstka and F. P. Stein, above note 58, p. 175; for a fuller discussion of the status of selfsynchronization, see B. J. A. van Bezooijen, P. J. M. D. Essens and A. L. W. Vogelaar, “Military Self-Synchronization: An Exploration of the Concept”, in *Proceedings of the 11th International Command and Control Research and Technology Symposium*, Cambridge, 26–28 September 2006, available at: www.dodccrp.org/events/11th_ICCRTS/html/papers/065.pdf.

程度。律师应该意识到，如果这种趋势按计划进行，则在一些事件的参与方面，将日益难以在各种自主系统之间加以区分。这种发展可能使审查新武器系统的过程以及确定系统开发或运行人员的责任变得更加复杂。事实上，试图在不同系统之间进行区分，例如致命的和非致命的系统之间的区别可能变得越来越人为。“多个系统所构成的系统”中的每个设备自身就是一个系统，会实施单独行为，而每个设备又是整个系统的一部分，因此还会表现出集体行为。

自主性和人道法

使用自主军事系统的法律影响

尽管武装力量的构成[63]以及冲突均将发生巨大变化，人们可能还是会问为什么将自主武器用于武装冲突的这一前景特别受法律关注。许多复杂程度各异的创新性机器对于门外汉来说是非常复杂的，这些机器用于战斗已有数千年，但是极少出现难以对其适用现有法律原则的情况。事实上，自主运作能力更为有限的武器系统已经在战斗中被使用，但对其合法性并不存在任何严重的争议。那么为什么一个自主性更高的武器会引起重大的法律问题呢？也就是说，为什么使用自主武器会影响国家履行国际人道法义务的能力，影响法院裁定是否违法的能力？

对这些问题的回答，自主系统与其他复杂军事系统之间在法律方面的重要区别在于，机器自主性影响到执行行动的决策过程，而其他复杂系统在决策作出后才会产生作用。[64]不过，正如其他地方所述，[65]这并不意味着系统具有足够自主性后可自行作出决策。相反，正如下面解释的那样，有两个关键的点：(1) 目前在武装冲突过程中作出的某些决定将从传统的决策者身上转移到界定自主系统行为及其使用负责人的身上；和 (2) 这样做，这些决策的性质将被改变。

63 Sharon Gaudin, “U.S. Military May Have 10 Robots per Soldier by 2023”, *Computerworld*, 14 November2013, available at: www.computerworld.com/s/article/9244060/U.S._military_may_have_10_robots_per_ soldier_by_2023.

64 人们认为这一点可能会引发进一步的问题，例如对于法律分析而言，“决策”的构成因素是什么，以及决策在满足何种条件时可被视为在自主系统的直接控制之内。

65 Armin Krishnan, *Killer Robots: Legality and Ethicality of Autonomous Weapons*, Ashgate, Farnham, 2009, p. 33.

自主性再次指定行动决策

自主机器首先运行软件，实施类似于动作执行的全部或部分决策过程，之后其他软件再引导机器执行该动作。换句话说，本来不得不决定是否执行某项行为的人将部分或完全不用作出该决定，或者本来可以发起或阻止某一行为的人将部分或全部不用从事该任务。相反，该决策的某些部分由程序员编写的软件有效执行，但由于此前程序员通常并不会对执行最终行动的决策产生有意义的贡献，因而会引发一些新的法律问题。例如，自主武器系统的设计是否必须让攻击规划人员实施的控制与他们对战斗员实施的控制相同？武器开发者或其他界定自主武器系统行为的人可能对违反国际人道法的行为负刑事责任吗？当法律规定人类为某职能负责时，如是否符合比例原则的分析职能，该职能的自动化在多大程度上是合法的？

使用自主水平较低的系统时，操作员只是不用进行本身不具备重大法律后果的低级别决策。例如，“自主导引/即发即弃”导弹在攻击目标过程中可以自主躲避拦截，并不会替按下“发射”按钮的操作员作出任何具有重大法律意义的决策。然而，当系统自主性较高时，机器的控制系统可以模拟作出更重要的决策。当某武器的自主ISR系统或瞄准系统在选择目标或提出发射方案方面 (例如，爱国者导弹系统的地面雷达) 起重要作用时，[66]武器操作员就不再进行攻击目标前的全部考量工作，并且不再是发射与否的唯一或主要决策者。在法律层面上，不能将机器本身视为决策者[67]，因此决策可部分或全部归因于自主系统行为负责人 (注意确定这些个人和组织本身可能就是一项复杂的工作)，以及负责决定使用武器的人。

自主性改变了决策性质

将操作决策从传统决策者转移到界定自主系统行为的人员身上，必然改变了这些决策的性质，体现在以下三个相互关联的方面。

66 Raytheon,“Patriot”, available at: www.raytheon.com/capabilities/products/patriot/.

67 也就是说，机器不是国际人道法的主体，而且人们认为它们今后变成主体的可能性较低。

首先，决策的普遍性必然会增加。当由人员在具体情况下“当场”作出的决策，被机器预设程序性指令所取代或补充时，那么具体情形下有关个别行为的决策就被涉及范围广泛的类似政策的多种选择所取代，该选择可适用于许多情形，前提是要符合机器预设的参数。

二，决策的时间变化。有关是否以及如何使用自主系统执行某一行动的决策，实际上是在相关行为被编写入机器程序以及使用武器时做出的，而不是在冲突局势中出现情况时。这一点很重要，因为首先它可能对国际人道法适用的时间范围有影响[68]，其次，它需要假定将来部署机器的情形在重要方面与机器开发和测试时设想的情形并没有什么不同。

第三，作出决定的信息基础被改变。通过自主系统实施的决定不能基于对决定所涉及的情况的直接 (甚至间接) 观察；相反，它们必须基于在机器编程时通过经验和远见而获得的任何信息，然后在决定使用武器时确认。

这些变化的一个结果是，当决策是部分或全部通过自主系统作出时，人类的某个具体决策和某个具体动作 (或重复动作) 或某个具体后果 (例如打击某个特定目标) 之间的因果关系就有可能被削弱。

无论动作是否被明确地编程到机器中，无论是否采用人工智能技术使机器自行灵活调整其行为，决策性质的上述变化都会发生；在这两种情况下，机器开发人员或其负责人都可以对系统的行为进行控制，方法就是确定目标并为系统配备实现该目标的一些手段。

确定法律问题

军事决策的这种改变是重要的，因为国际人道法也与决策有关。国际人道法力图“出于人道原因，限制武装冲突的影响”[69]，试图引导将个人和冲突各方的决策以及最后采取的行动。在武装冲突中，当自主技术实施的决策

68 For a discussion of one such question, see Tim McFarland and Tim McCormack, “Mind the Gap: Can Developers of Autonomous Weapons Systems be Liable for War Crimes?”, U.S. Naval War College International Law Studies, Vol. 90, 2014, p. 361, available at: www.usnwc.edu/getattachment/ed8e80adb622-4fad-9a36-9bedd71afebe/Mind-the-Gap–Can-Developers-of-Autonomous-Weapons.aspx.

69 ICRC, “War and International Humanitarian Law”, Geneva, 2010, available at: www.icrc.org/eng/warand-law/overview-war-and-law.htm.

受国际人道法规制时，最有可能出现机器自主性的使用与法律不符的情况；也就是说，这时使用具有自主运行能力的武器或其他系统改变了某一决策过程或决策结果，超出国际人道法允许的范围或以其他方式妨碍法律的运作。鉴于当今武器系统中存在的自主能力水平较低，这种事件仍然是假设的，但似乎可以假设：如果自主ISR系统直接向自主武器系统提供信息，以致人类指挥官无法在激活武器之前独立验证其准确性，可能被视为违反了在攻击中采取预防措施的义务。

这似乎是一个模糊的标准，但鉴于自主系统的形式和应用的预期范围以及当前的发展初期阶段，也许是最可靠的说法。但是，以下准则可能有助于确定法律问题。

人类仍是法律决策者

国际人道法的本质和规定都是以人为中心的。其主要目的之一是“抑制战争的邪恶，减少不必要的苦难并改善战场伤员的命运”[70]，因此其重点必然在于人，即武器的使用者以及受武器影响的人。法律发展的推动力，是规制并解释人的决策和行动的必要性，以及减轻不必要人类苦难的必要性。

这或许是老生常谈，但它有助于讨论该问题。一项新技术的发展，特别是许多人认为是“军事革命”先兆的技术发展[71]，自然引起人们对技术细节及其带来的可能性的关注。人们容易将技术 (或使用技术的装置) 作为法律的焦点，而不关注人。在某些情况下这是适当的；当某一技术或装置按照预期将对敌对行动和所涉人员产生有限且和明确的影响，认为该技术或装置体现了对法律的某种影响，这是有利的。旨在使人永久失明的激光武器具有有限且和明确的效果，可以容易地分析出它是否符合武装冲突法，很容易看出这种激光武器必然违反了武装冲突法的限制规定；因此，律师重点关注这些设备带来的某些不利影响，并且对其进行相应规制，这是适当的。[72]但是，自主系统并非如此。

70 《改善战地武装部队伤者境遇之日内瓦公约》，1864年8月22日（1865年6月12日生效），序言。

71 P. W. Singer, above note 54, p. 203.

72 《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》的《第四附加议定书（激光致盲武器议定书）》，1380 UNTS 370，1995年10月13日（1988年7月30日生效）。

现有和新型军事系统新增加的自主能力，带来了不同的挑战。一个系统某种程度的自主运作能力所产生的效果并不容易确定，不容易根据国际人道法的要求对其效果进行评估。如上所述，自主控制系统仅直接影响武器和操作者之间的关系，从而直接影响操作者的行为，而不一定影响被控制武器的行为。另外，考虑到时间上以及可归为机器自主性的功能上的可变性，很难就操作员或开发者在战场上特定事件中的行为 (思想和行为) 做出一般性的陈述。可以确定的是，律师在确定何处会出现法律问题时，必须关注该行为。赋予机器自主运作能力的技术只是有助于作出这些判定。

区分技术问题和法律问题

法律学者最常提出的一些问题基本上是关于拟议的自主系统是否满足遵守法律所需的某些性能标准。这种问题目前往往与区分原则和比例原则最为相关。[73]

例如，就区分原则而言，有些作者认为，现有机器人系统无法可靠地区分战斗员和平民，原因包括传感器系统的有效性不足[74]，控制器软件无法充分“理解” 环境和人类行为，难以在复杂情况下区分有关和无关信息[75]。还有类似的观点认为，机器人不能判断所提出的攻击是否符合比例性要求，原因是该问题涉及定性问题并具有主观性质。[76]

其他作者提出机器人系统在冲突中可靠运作的能力具有一般局限性。Asaro认为自主系统“学习与适应能力极其有限”是“难以或无法设计出能够应付战争迷雾和摩擦的系统”的原因。[77]其他人指出，复杂的自主系统的行为是难以可靠地预测出来的。[78]

---

73 See, e.g., C. Grut, above note 37; Peter Asaro, “On Banning Autonomous Weapon Systems: Human Rights, Automation, and the Dehumanization of Lethal Decision-Making”, *International Review of the Red Cross*, Vol. 94, No. 886, 2012, p. 687; M. Wagner, above note 16.

74 See, e.g., Benjamin Kastan, “Autonomous Weapons Systems: A Coming Legal ‘Singularity’?”, *Journal of Law*, *Technology and Policy,* No. 1, 2013, p. 60.

75 M. S. Riza, above note 9, pp. 129–132.

76 See, e.g., Noel E. Sharkey, “The Evitability of Autonomous Robot Warfare”, *International Review of the Red Cross*, Vol. 94, No. 886, 2012, p. 789.

77 P. Asaro, above note 73, p. 692.

78 A. Finn and S. Scheding, above note 49, p. 36: “As the degree of autonomy increases, so it becomes increasingly difficult to predict the sum state of the system.” Also see the discussion of developer accountability on p. 183.

考虑到当今相关技术的发展状况，上述许多观点无疑都有可取之处。然而，如果对使用自主武器持有反对意见，是为了说明此类系统在特定情形中按照国际人道法要求的标准进行运作的能力，那么即使该说法是准确的，它也仅涉及机器能力的某些具体方面，且仅对某个时间点有效；随着技术的发展，可能就不存在这些反对意见了。因此，律师应避免根据这种观点就自主系统的合法性发表一般性的陈述。只有预期此类反对意见会持续存在的情况下，如存在某个绝对会一直阻碍具有自主运行能力的系统遵守国际人道法的因素，或机器自主性本身固有的特征会违反国际人道法，此类反对意见才能成为法律裁决的基础 (或许可以指导国家武器审查进程)。机器自主性范围之外的例子包括禁止背信弃义的行为 (由于行为的性质) 和禁止使用沙林毒气 (由于其固有的滥杀滥伤作用)。如果不确定存在此类因素，那么仅说明某一特定系统在某个特定情况下是否可合法使用，似乎才更恰当。因此，建议对自主系统合法性的调查应包括一个资质审查问题：如果系统可以，甚至即使在假设的情况下可以，在某个情况下与非自主系统 (无论是人工操作还是其他方式) 一样有效运作，那么是否还有法律依据反对其使用？如果答案明显为“否”，那么该反对意见则更应被视为一种技术挑战，克服挑战之后，该系统才能被合法使用，或者被视为限制系统使用情况的理由，直到克服该挑战为止。例如，如果一个自主瞄准系统能够在不受协助的情况下与人类战斗员一样可靠地，或者比人类战斗员更可靠地，在战斗员和平民之间加以区分，那么是否应反对其使用？这表明由于难以预测自主系统在所有情况下的行为或激活后存在计算机错误和故障的可能性而对自主系统使用持反对意见的情况，也将属于这一类。尤其是，出现故障并不是自主系统所特有的;而是所有依赖计算机执行部分任务的任何军事系统都必须解决的问题。

然而，如果答案为“是” (即不管自主系统的有效性如何，其使用目前不符合国际人道法的限制规定)，那么律师需要更谨慎地确定应如何限制或引导自主系统的发展，或应如何制定法律。例如，如果发现法律要求对每一次单独的攻击或暴力行为进行一定程度的直接人类控制，情况就可能如此。如果是这样，无论系统的运行情况如何，都应不允许将攻击完全置于自主系统的控制之下。

结论

尽管自主系统仍处于早期发展阶段，但具有足够自主性的军事系统的使用似乎也可能会在一系列基本方面检验国际人道法的极限。自主系统的新颖性在于它们重新分配操作决策，并由此改变了这些决策的性质。在武装冲突的情况下，个别战斗员根据情况作出的决策将被界定自主系统行为的人作出的一般性选择所取代。

预期拟议系统中自主运行程度所具有的可变性，自主系统的潜在应用范围以及自主系统相互之间直接协作的可能性，这三者结合在一起导致侧重于具体情况下特定机器行为的分析方法更加复杂。相反，建议律师将机器能力视为武器开发者所作出决策的结果，并关注开发人员以及部署和使用自主系统的操作人员所发挥的作用。尤其是，由于自主能力而新产生的法律后果与机器执行其功能的好坏无关，仅与其开发人员和操作人员如何参与结果有关。